ESG-constrained portfolio optimization

Thesis project for
MSc Quantitative Finance and Actuarial Science

Youssef Ben Jemia
(2080053)

Tilburg School of Economics and Management
Tilburg University

Abstract

This project builds upon the existing litterature about sustainable investing and ESG integration into portfolio optimization frameworks. We consider two metrics to evaluate a company's performance: the carbon footprint measured by carbon intensity and an estimated carbon risk exposure using the Fama-French 3 factor model augmented a BMG carbon factor. We use two extensions of the traditional mean variance framework, including the ESG-efficient frontier developped by (Pedersen et al., 2021) and the benchmark optimization framework. Following the work of (Gorgen et al., 2020) we construct a BMG carbon factor using carbon intensity as the sorting measure. We use a sample of 1899 stocks from the STOXX Europe TM index as test assets from July 2007 until June 2022. We found no proof of a carbon risk premium during the sample period, however the BMG factor provided some complementarity to the traditional factors including the size, value and momentum factors. Finally, we find, using the (Pedersen et al., 2021) framework that ESG-motivated investors experienced lower excess returns from 2016 until 2020 compared to investors maximizing utility solenly from risk-return. More importantly, we find that this trend is reversed post 2020, suggesting potential better risk-adjusted returns for sustainable investing strategies in the future.

Acknowledgements

I would like to thank Prof. Dr. Anja De Waegenaere for her supervision in the writing of the thesis. As a student, I still have plenty of knowledge to learn. But discussing econometrics with her was an invaluable source of learning.
I would also like to thank my family for their unconditional support throughout this journey, je vous aime.

Contents

1 Introduction

Climate change represents the greatest challenge facing humanity in the 21st century! From the World Economic Forum to major accord like the Paris Agreement, governments along with major financial institutions have put forward a clear vision for collective action to fight this crisis. Consequently, Sustainable Investing has been gaining momentum in a context where mindsets towards Environment, Society and Governance, that is to say the integration of sustainability considerations into investment processes, have drastically evolved.

ESG used to be an ethical box to check. It is now impossible to miss. Investors are facing dire pressure from across the board to integrate sustainability at the core of their investment process. From regulators, with the European Union spearheading ESG standardisation initiatives. From clients, who want sustainable portfolios but are (rightly) confused by 'ESG' products. And by civil society at last, with rising concerns about Climate Change impact on business operations.

This urgency has made sustainable investing essential for channeling capital towards climate solutions as governments and companies prioritize shifting to a greener economy. Reflecting this, ESG-focused investments have soared and are expected to continue apace in the near future. In aggregate, the share of ESG assets over total Assets under managements (AuM) are expected to increase from 14.4% ($18.4tn) in 2021 to 21.5% in 2026 overall, accumulating over $34tn in value[1]. Within Europe, this growth is expected to reach 53% by 2026, accumulating over $19.5tn according to a study shared by PWC[1].

This massive inflow of capital means incorporating ESG factors into portfolio construction is increasingly important, both for sustainability concerns and risk mitigation. As the economy transitions towards greener technologies, portfolios require resilience against stranded asset and transition risks. Thus optimizing for ESG alongside financial objectives allows investors to address climate change while navigating the transition shifts underway.

In light of the above, this thesis project addresses common quantitative methods to integrate ESG criterias into commonly used portfolio allocation problems. We propose two extension of the traditional Mean Variance optimization introduced by (Markowitz, 1952). Additionally, we define the Greenhouse Gas (GHG) emissions intensity to assess the ESG performance of a company. We show two alternatives for using this carbon data, first as an additional variable in optimization problems then as a metric for financial risk evaluation. The latter involves estimating the 'carbon risk' by means of factor models and follows the work of (Gorgen et al., 2019). Finally, we compare the performance of carbon efficient and less efficient companies since 2007 in terms of risk-adjusted returns, for this we also estimate the risk premium associated with investing into carbon efficient companies. Hence, the key research questions are: What are the most effective quantitative methods for incorporating carbon emissions data into portfolio construction and asset allocation? Does carbon risk carry a risk premium? Does explicitly managing carbon risk lead to improved investment outcomes compared to traditional approaches that ignore ESG factors?

[1] https://www.pwc.com/gx/en/financial-services/assets/pdf/pwc-awm-revolution-2022.pdf

2 Litterature review

The key idea behind integrating ESG related data into portfolio allocation is not new and has spurred the attention of several research for over a decade. However, different research use different approaches to integrate ESG data into portfolio selection. Bilbao et al. (2013) used a two-step model, optimizing expected wealth before calculating a social satisfaction index to find the best financial-social portfolio. Other papers such as (Hirschberger et al., 2013; Utz et al., 2015) add sustainability as a third optimization criterion alongside risk and return. Chen et al. (2021) proposed a three-step method using data envelopment analysis, an ESG-filtered universe, and standard optimization. This research builds on the methodology of (Pedersen et al., 2021), who incorporated relative ESG scores into the objective function to examine their impact on optimal allocations. They introduced an ESG utility function into the traditional mean variance utility and derived optimal allocations for investors while controling for sustainability preferences. Their approach yielded four-fund separation portfolios consisting of the risk-free asset, tangency portfolio, minimum-variance portfolio, and ESG-tangency portfolio.
Additionally, recent research has focused on incorporating ESG related data into asset pricing models to estimate the financial risks associated with poor ESG performance. This idea was first proposed by (Gorgen et al., 2019). Their work extended the Fama-French-Carhart model by including a brown-minus-green (or BMG) risk factor. Using the sorted portfolios technique popularized by (Fama and French, 1992), they build a factor-mimicking portfolio based on a scoring model and more than fifty carbon risk variables. They then defined the carbon financial risk of a stock using its price sensitivity to the BMG factor or its carbon beta. Since the release of this paper, further research followed this methodology by using alternative proxies for the ESG data. This include the work of (Roncailli et al., 2020) which compared the BGS composite of (Gorgen et al., 2019) with a number of alternative measures. They concluded that the carbon intensity, is the most correlated proxy to the BGS composite, which rises the motivation of using this metric as an ESG metric. However, results to whether the carbon risk is priced in the market are mitigated. Gorgen et al., (2019) and (Roncailli et al., 2020) found negative results, whereas (Bolton and Kacperczyk, 2021; Bolton and Kacperczyk 2022; and Hsu et al., 2023) find that climate risk are priced, when proxied by carbon emissions.
. Finally, there has been extensive research on whether sustainable investment strategies outperform traditional investment strategies in terms of long-term financial performance. Some studies suggest that there is no significant difference in financial performance between ESG-focused and traditional investments (Renneboog et al., 2008; Halbritter and Dorfleitner, 2015). However, other studies suggest that SRI funds can achieve superior returns. This is supported by a large majority of empirical studies including the meta analysis conducted by (Clark et al., 2014) who concluded that "80% of the studies show that stock price performance is positively influenced by good sustainability practices".

This thesis is constructed as follows: We start by defining common concepts in sustainable investing including the materiality of ESG data in section 3. In section 4 we introduce the ESG risk measure including the methodology for estimating and validating the financial risk associated with carbon emissions, including the major statistical tests. Following on this, we introduce the theoretical background for solving the two extended versions of the Mean Variance optimization framework. Finally, empirical results are generated in Section 6.

3 Sustainable investing in a nutshell

The European Commission (EC) defines the concept of sustainable finance as follows[2]:

> "*Sustainable finance refers to the process of taking* ***Environmental, Social and Governance (ESG)*** *considerations into account when making investment decisions in the financial sector, leading to more long-term investments in sustainable economic activities and projects.* ***Environmental considerations*** *might include climate change mitigation and adaptation, as well as the environment more broadly, for instance the preservation of biodiversity, pollution prevention and the circular economy...*"

Within this definition, the EC introduces the concept of ESG (Environmental, Social, and Governance) that has gained popularity among asset owners and managers during the last two decades. However, the line between sustainable finance (SF) and ESG remains unclear. The same confusion extends to other frequently used terms like impact investing, sustainable investing (SI), and socially responsible investing (SRI). Though these terms differ, they share an underlying idea, in evaluating and integrating ESG factors into investment analysis.[3] For this project, we will refer to these concepts broadly as "sustainable investing".

3.1 ESG materiality: defining what is "good"

An essential concept in the realm of ESG is **materiality**. The concept of materiality is critical for sustainable investing, as it helps identify and prioritize the ESG factors most relevant to a company's financial performance and long-term value.
However, material ESG factors differ across industries and companies. For carbon-intensive sectors like oil and gas or utilities, climate change risks including greenhouse gas (GHG) emissions, energy transition, and shifting regulatory requirements are often highly material. For example, an energy company that emits a lot of GHGs may be subject to stricter regulations, which could increase its costs. In contrast, for technology companies, data privacy, information security, and AI ethics may be more material ESG issues. Thus, assessing ESG materiality allows companies to focus sustainability efforts on the most significant risks and opportunities. From an investment perspective, analyzing ESG materiality enables investors to identify companies that are strategically managing the sustainability risks most relevant to their business. However, a lack of reporting standards and common definition of ESG risks has impeded the usefulness of ESG information for investors, by reducing the comparability and reliability of such data (Amel-Zadeh et al., 2018; Kotsantonis et al., 2019).
As a result, a number of voluntary and mandated ESG disclosure frameworks have emerged aimed at improving sustainability reporting and providing investors with decision-useful climate and ESG data. Among the most widely adopted voluntary frameworks are the **Partnership for Carbon Accounting Financials (PCAF)** and the **Task Force on Climate-Related Financial Disclosures (TCFD)** which provides a disclosure framework for companies to report on climate-related financial risks and opportunities across governance, strategy, risk management, and climate metrics/targets.[4] More importantly, it provides a framework for financial institutions to report on the climate and ESG performance associated with a portfolio of investments. These will be used further in this thesis to quantify such performances.

[2] https://finance.ec.europa.eu/sustainable-finance/overview-sustainable-finance_en#what
[3] https://www.unpri.org/download?ac=4571
[4] https://www.fsb-tcfd.org/example-disclosures/

3.2 Sustainable investment strategies

As mentioned above, sustainable investing has become increasingly popular among investors and fund managers. However, different institutions use different strategies to incorporate ESG information into asset allocation problems (Amel-Zadeh et al., 2018). Particularly, different reporting frameworks use different labels to categorize these strategies. For example, the UN PRI (Principles for Responsible Investment) framework[5] and Eurosif[6] define sustainable investment strategies under 7 categories which are shown in Table [1]. Below we provide some example of how these are being used in the asset management industry.

Eurosif	PRI
Exclusion of holdings from investment universe	Negative/ exclusionary screening
Norms-based screening	Norms-based screening
Best-in-Class investment selection	Positive best-in-class screening
Sustainability themed investment	Sustainability themed investing
ESG integration	Integration of ESG issues
Engagement and voting	Active ownership and engagement
Impact investing	-

Table 1: Definition of the different sustainable stratgies under the Eurosif and the PRI frameworks.
Source: EUROPEAN SRI STUDY 2018 report, available at: https://www.eurosif.org/wp-content/uploads/2021/10/European-SRI-2018-Study.pdf

3.2.1 Screening strategies

The UN PRI defines screening strategies as "the use of a set of filters to determine which companies, sectors or activities are eligible or ineligible to be included in a specific portfolio".[7] These filters in question are chosen differently among different investors. For example, Arabesque Asset Management which is one of the first asset management companies that is fully specialised in sustainable investing, uses screening to define their investment universe:[8] they only inlude stocks of companies with the top 75% ESG scores (Positive screening) and exclude stocks of companies that do not comply to the UN Global Compact compliance (Negative screening). Other filters might be based on sector preference, levels of GHG emissions, or climate-related risks.[9]

3.2.2 ESG integration

ESG integration strategies have become more popular in recent years, thanks to advances in computation systems and the availability of ESG data. These strategies consider a broader set of ESG indicators than screening strategies, which have been popular for many decades.
This involves explicitly using ESG factors alongside financial factors to assess how ESG issues can impact a company's financial performance, both positively and negatively.

[5] https://www.unpri.org/Uploads/i/m/n/maindefinitionstoprireportingframework_127272_949397.pdf

[6] https://www.eurosif.org/responsible-investment-strategies/

[7] https://www.unpri.org/introductory-guides-to-responsible-investment/an-introduction-to-responsible-investment-screening/5834.article

[8] UN PRI 2023 reporting available at: https://www.icgn.org/sites/default/files/2021-08/PRI_apracticalguidetoesgintegrationforequityinvesting.pdf

[9] Eurosif 2023 report available at: https://www.eurosif.org/wp-content/uploads/2023/05/Eurosif-Report-on-Climate-related-Data.pdf

Examples include but are not limited to, using ESG metrics as an extra input parameter in traditional portfolio optimization framework (Pedersen et al., 2021; Cheema-Fox et al., 2021). In practice, investors are not limited to one SI strategy and can incorporate elements of both ESG integration and screening strategies.

3.2.3 Other strategies

Other strategies in Table [1] include norms-based screening based on compliance with international norms, and sustainability themed investment, defined by Eurosif as "Investment in themes or assets linked to sustainability development" with a focus on companies related to health, green energy or sustainable agriculture.
In this context, Eurosif conducted a survey to summarize the main strategies used by European asset managers to incorporate ESG related data into investment decisions. For this, they conducted a survey of 37 EU financial industry practitioners, including asset managers and pension funds, to gain insights into the latest sustainable investing strategies in the EU. Figure [1] summarizes the most common practices among EU asset managers for incorporating ESG data into investment strategies.

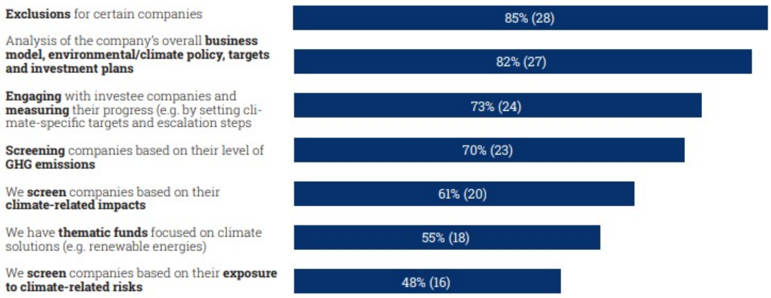


Figure 1: Most common practices in using climate related data by EU asset managers and pension funds in 2022, according to a survey conducted by Eurosif over 37 practicioners.

These include the use of the GHG emissions and climate related risks which will be the focus of this thesis project. Specifically, we will carbon emissions data along traditional portfolio methods to draw optimal investment strategies that depends on the investor's preference for ESG performance. Thus, we will focus on a narrower frame of sustainable investing which excludes the Social (S) and Governance (G) pillars of ESG to focus on the Environmental (E) pillar.

4 ESG risk metrics

In order to introduce ESG information into traditional portfolio optimization, we first introduce a risk metric to quantify the performance of financial assets with respect to such an ESG metric. In this subsection we introduce the notion of ESG risk measure. Specifically, we introduce two relevant metrics: The carbon footprint and the carbon risk. The latter measure uses traditional factor models to estimate the financial risk that stems from poor carbon emissions. For this we also introduce the necessary steps to validate such models in Section (4.3).

Denote by $\mathcal{S}_i$ an ESG risk metric associated with an asset i. Following (Le Guenedal et al., 2022) we assume that this metric is additive and that the portfolio ESG score is the weighted average of the individual scores. For a vector of portfolio weights $\boldsymbol{\pi} = [\pi_1, \pi_2, ..., \pi_n]'$, we have:

$$\mathcal{S}(\pi) = \sum_{i=1}^{n} \pi_i \mathcal{S}_i \tag{1}$$

where $\mathcal{S}(\pi)$ is the portfolio ESG score.
There is a substantial number of metrics that can be used to evaluate the performance of a company's ESG performance, including ESG ratings from external rating agencies such as MSCI, Sustainalytics or Refinitiv. In their paper, (Le Guenedal et al., 2022) summarizes the various climate risk measures that are used in the asset management industry. However, we will make use of the Greenhouse Gas (GHG) emissions as a prominent measure to categorize companies into sustainable companies and less sustainable ones. A natural advantage of using carbon emissions is that these are company reported data, thus they can be compared accross all companies and are less prone to inconsistency between different data providers (Le Guenedal et al., 2022).

4.1 Carbon footprint

The carbon footprint which quantifies the pollution of a company in terms of its carbon intensity i.e. its Greenhouse Gas (GHG) emissions normalized by its revenues:

$$\mathcal{S}_i = \frac{\mathcal{CE}_i}{\text{Revenues}_i} \tag{2}$$

where $\mathcal{CE}_i$ denotes the carbon (GHG) emissions associated with asset i.
We use the framework proposed by the TCFD to allocate a carbon footprint score at the portfolio level. This can be done using the Weighted Average Carbon Intensity (WACI) metric.[10] For a portfolio allocation $\boldsymbol{\pi} = [\pi_1, \pi_2, ..., \pi_n]'$, the portfolio ESG score using the carbon footprint is defined as:

$$\begin{aligned} \mathcal{S}(\pi) &= \text{WACI}(\pi) && (3) \\ &= \sum_{i=1}^{n} \pi_i \cdot \frac{\mathcal{CE}_i}{\text{Revenues}_i} && (4) \\ &= \sum_{i=1}^{n} \pi_i \cdot \mathcal{CI}_i && (5) \end{aligned}$$

where $\mathcal{CI}_i$ denotes the carbon intensity associated with asset i and is used to track companies that are most efficient in managing their carbon emissions i.e. companies that are able to produce extra revenues while polluting the least.

[10] https://assets.bbhub.io/company/sites/60/2020/10/FINAL-TCFD-Annex-Amended-121517.pdf

4.2 Carbon risk

The second measure of interest is a carbon risk defined as a company's exposure to transition risks i.e. the risks associated with a company when shifting to a greener economy. The interest in the carbon exposure arises from the fact that, as the regulation over carbon emissions are tightening in order to reduce the overall CO2 emissions (examples include carbon pricing[11]), firms currently with high carbon emissions will be penalized in the future in comparison with firms currently with low carbon footprints (Gorgen et al., 2019). Should this theory hold, the carbon risk exposure would provide a forward-looking measure of a company's financial risks associated with a poor carbon management. However, in contrast to the carbon footprint which stems from fundamental data that is reported by the companies, carbon risk is estimated using factor models.

The idea of estimating a carbon risk using factor models, was first proposed by (Gorgen et al., 2019) who build a brown minus green (BMG) carbon risk factor and use it as an extra risk factor in the Carhart 4 factor model (Carhart, 1997).

They first define a Brown-Green-Score, using over 50 carbon emissions and environmental metrics, then the BMG factor is constructed using the methodology of sorted portfolios by (Fama and French, 1992).

Due to data availability, we cannot use the metrics defined by (Gorgen et al., 2019) to define the BGS score, however (Roncalli et al., 2020) showed that the composite indicator built by (Görgen et al., 2020) is well captured by a factor based on the GHG emissions intensity only. We therefore consider GHG emissions intensity ($\mathcal{CI}$) as a robust metric for identifying companies most exposed to transition risks.

In addition, following the work of (Loyson et al., 2023; Roncailli et al., 2020) we estimate the carbon risk by using the Fama-French 3 factor model (Fama and French, 1992) augmented by the carbon factor (BMG). This is equivalent to estimate the following regression:

$$r_{i,t} - r_f = \alpha_i + \beta_i^{MKt}\left(R_t^{MKt} - r_f\right) + \beta_i^{SMB} R_t^{SMB} + \beta_i^{HML} R_t^{HML} + \beta_i^{BMG} R_t^{BMG} + \epsilon_{i,t} \tag{6}$$

where $r_{i,t} - r_f$, $(R_{i,t}^{Mkt} - r_f)$ denote the asset i's and the market excess returns, R_t^{HML},R_t^{SMB},R_t^{BMG} are the returns of the size factor, the value factor and the BMG factor. β_i^{Mkt},β_i^{SMB},β_i^{HML},β_i^{BMG} represent the asset i's exposure to the market risk and the aforementioned factors. $\epsilon_{i,t}$ represent the idiosyncratic error terms which are assumed to be identically and independently distributed, with mean zero.

Equation (6) can be estimated using a multivariate OLS regression by regressing the asset excess returns, on the factor returns. In this case, $\hat{\alpha}_i$ can be thought of as the average proportion of excess return that is not explained by the exposure to the aforementioned factors. Thus, factor models that better explain the returns $r_{i,t}$, will tend to generate models with a smaller $\hat{\alpha}_i$.

We can then define the ESG measure associated with the carbon risk as:

$$\mathcal{S}_i = \hat{\beta}_i^{\text{BMG}} \tag{7}$$

where the portfolio carbon risk is:

$$\mathcal{S}(\pi) = \sum_{i=1}^{n} \pi_i \cdot \hat{\beta}_i^{\text{BMG}} \tag{8}$$

In addition, since a higher level of CO2 intensity is translated as a higher exposure risk, a negative coefficient for $\hat{\beta}^{BMG}$ implies that the asset i will be positively impacted by the transition

[11] https://www.worldbank.org/en/programs/pricing-carbon

to a greener economy, while a positive level is translated as negatively exposed to transition risks (Gorgen et al., 2019).
Although factor models have received a great attention in the Finance litterature, a number of assumptions have to be validated in order to obtain unbiased estimators of the risk exposures. First, by definition the linear factor models assume a linear relationship between the asset returns and the factors, thus nonlinear relationship cannot be estimated accurately using these mdoels. Additionally, we assume the error terms to be independently and identically distributed with mean zero. This implies (Brooks, 2014):

$$\text{Var}(\epsilon_{i,t}) = \sigma^2 \quad \forall i \tag{9}$$

$$\text{Cov}(\epsilon_i, \epsilon_j) = 0 \quad \text{for } i \neq j \tag{10}$$

$$\text{Cov}(\epsilon_t, R_t^k) = 0 \tag{11}$$

These assumptions are also referred to as homoskedasticity, no correlation and no multicolinearity. In simpler terms, we should assume that the returns of each variable are stationary and that the factors are orthogonal to each other, meaning they individually provide valuable information that can explain the cross-section of asset returns.
Furthermore, adding a new factor to the traditional factors such as size and value requires the validation of extra assumptions. According to (Görgen et al., 2020), there are two main criterias for a new factor to be eligible in the factor models: 1) generating extra performance or reducing risk compared to traditional factors, and 2) being complementary to traditional factors.
In what follow we present the methodology to construct and validate the augmented Fama French factor model. Specifically we explain the methodology of the sorted portfolios, necessary to build factor models and explain and the necessary tests to validate the BMG factor.

4.3 Building and testing the BMG factor

4.3.1 Fama French sorted portfolios

In order to implement factors into regressions of the form (6), factor models relies on the construction of factor mimicking portfolios (FMP) popularized by (Fama and French, 1992). These involve, creating portfolios by sorting assets each year, according to a characteristic that is supposed to explain the cross-variation of returns.
The process starts each June of year t by filtering the stock universe to keep only those with sufficient and reliable data for the specified factors. Typically, this involves requiring an available market capitalization in June of year t and an available variable Y in December of year t-1, where Y represents the factor-specific data needed. For example, Y could be book-to-market equity if constructing the size (SMB) factor portfolio or carbon intensity for the BMG carbon factor. Having defined the set of asset that will be used in the sorting of year t, these are then sorted into 2 size groups based on market capitalization of June of year t. To do this, we should specify a breakpoint value that will be used to sort asset into different groups. In this case we will use the median breakpoint. Thus, the first sorting will generate 2 portfolios, namely big (B) and small (S) assets. Assets are also sorted based on the variable Y into 3 groups. For this, we also define 2 breakpoints. We will use the 30th and 70th percentiles as breakpoints. This yield 3 portfolios sorted on the Y variable. The intersection of the 2 size groups and the 3 groups sorted on the Y variable will result in 6 portfolios as shown in table [2].

	Low Y	Medium Y	High Y
Small	-	-	-
Big	-	-	-

Table 2: Illustration of the 3x2 factor portfolios formed on size and a variable Y

Returns of these 6 portfolios are calculated using either an equally weighted scheme or a value weighted scheme. In our case we use the value weighted scheme. Thus, weights of assets in each group are calculated each year, by dividing the asset's i market cap, which has been used for the sorting, by the sum of the market caps in the portfolio it belongs to. This generates a set of 6 portfolio returns which are held for one year from July of year t until June t+1. Since portfolios are held for 1 year, assets should also display available returns from July t until June $t+1$ to be included in the sorting of year t. Additionally, each year we winsorize the data at the 1% and the 99% based on the Y variables. This is done to limit the effects of outliers in estimation. This process is repeated each year in June, using newly available data, including companies that started to report the necessary data and recalculating the new breakpoints. Finally this yield a time series of the 6 portfolios from July 2007 until June 2022. The factor return is then constructed as the difference between the average returns of portfolios. Further details will be explained in each factor construction.

Size and value factors

Size and value factors are constructed as the intersection of groups based on size and the book-to-market equity (BE/ME) variable. Following (Schmidt et al., 2019) we define the BE/ME as:

$$\text{BE/ME} = \frac{1}{\text{PBVR}} \tag{12}$$

where PBVR denotes price/book value ratio. Additionally, we remove negative values of the BE/ME. Thus for an asset to be included in the sorting of year t, it must have available size (mc) data in June of year t, available returns from June t until July t+1 as well as available and positive BE/ME as of December t-1. 6 portfolios are constructed based on the intersection of 2 size groups and 3 BE/ME groups namely, Growth (G), Neutral (N) and Value (V) i.e.

	Growth (G)	Neutral (N)	Value (V)
Small (size)	SG	SN	SV
Big (size)	BG	BN	BV

Table 3: Illustration of the 3x2 factor portfolios formed on size and BE/ME

The returns for the size factor (SMB) and the value factor (HML) are then constructed as follows:

$$\text{SMB}_t = \frac{1}{3}\Big(\text{SG}_t + \text{SN}_t + \text{SV}_t\Big) - \frac{1}{3}\Big(\text{BG}_t + \text{BN}_t + \text{BV}_t\Big) \tag{13}$$

$$\text{HML}_t = \frac{1}{2}\Big(\text{SV}_t + \text{BG}_t\Big) - \frac{1}{2}\Big(\text{SG}_t + \text{BG}_t\Big) \tag{14}$$

Where X_t denotes the return of portfolio X in t.

BMG factor

Following the methodology of (Gorgen et al., 2019), we can construct the BMG factor by using the carbon intensity metric as the sorting variable. In the second sorting we form 3 groups based on the terciles breakpoints. Assets in the bottom tercile are then Green (G) assets since they are the most efficient companies, while companies in the top tercile are brown firms (B). 6 portfolios are then constructed as follows:

	Green	Neutral	Brown
Small	SG	SN	SB
Big	BG	BN	BB

Table 4: Illustration of the 3x2 factor portfolios formed on size and carbon intensity

Returns on the BMG factor are then constructed as (Gorgen et al., 2019):

$$\text{BMG}_t = \frac{1}{2}\Big(\text{SB}_t + \text{BB}_t\Big) - \frac{1}{2}\Big(\text{SG}_t + \text{BG}_t\Big) \tag{15}$$

Momentum factor

We also construct the momentum factor proposed by the Carhart 4 factors models (Carhart, 1997). is constructed using a slightly different methodology than the other factors. The portfolios are not constructed every June of year t and held for 1 year, instead these are constructed every month τ. First we need to define the momentum measure. The momentum measure for month τ is defined as the average returns over $\tau - 11$ until $\tau - 1$. Furthermore the size group is also sorted every month using the lagged month $\tau - 1$ market cap. Thus to be included in the sorting of month τ, an asset must have an available size value and an available momentum value for $\tau - 1$ (available returns from $\tau - 12$ until $\tau - 2$). Portfolios corresponding to the month τ is then the intersection of 2 size groups and 3 groups based on momentum, namely Losers (L), Medium (M) and Winners (W):

	Losers	Medium	Winners
Small	SL	SM	SW
Big	BL	BM	BW

Table 5: Illustration of the 3x2 factor portfolios formed on size and Momentum

Returns for the momentum factor is then calculated as:

$$\text{WML}_t = \frac{1}{2}\Big(\text{SW}_t + \text{BW}_t\Big) - \frac{1}{2}\Big(\text{SL}_t + \text{BL}_t\Big) \tag{16}$$

Market factor

We also construct value weighted for the market factor. This factor is also constructed monthly. For each sorting month τ, we include every company that provides returns data for τ and market cap data for $\tau - 1$ and weight them according to their market cap.

Having defined the methodology for constructing the factor portfolios, we can postulate postulate the necessary tests to validate the model specified in (6). First, we present the tests for validating the homoskedasticity, no correlation and no multicolinearity assumptions, before postualiting the tests to validate the BMG factor as a new factor.

Validating linear factor models

Recall that, linear models are built upon certain assumptions, including homoskedasticity, no correlation and no multicolinearity. Multicollinearity means some of the regressors correlate with each other. It is problematic if there is a high correlation between the regressors because, the explanatory power of the model will be high, but individual variables will not be significant (Brooks, 2014). This can be tested by plotting the correlation matrix of the underlying factors, or more formally by measuring the Variance Inflation Factor (VIF) after estimating (6). This can be done using the VIF() command from the **regclass** package in R. No formal criteria exist for deciding when a VIF is too large, but generic cutoff values of <5 and <10 are commonly used (Craney and Surles, 2002).
Additionally, we can test for the heteroskedasticity and collinearity of the residuals. The former assumption implies that the variance of the error terms are constant over time, thus after estimating the model (6), we can plot the residuals over time and check whether a structural change has occured in the data. Additionally, collinearity can be tested using the Durbin and Watson test. This builds upon the following hypothesis testing:

- $\mathbf{H}_0$: There is no correlation among the residuals.
- $\mathbf{H}_\text{a}$: The residuals are autocorrelated.

Durbin and Watson test can be done using the **car** package in R. In case these assumption is violated, further steps can be taken to improve the statistical efficiency of the model. This include using alternative regression models such as the Generalized least squares regressions or the considering robust standard errors such as the heteroskedasticity and autocorrelation consistent (HAC) standard errors (Brooks, 2014).

Complementarity of the carbon risk factor

To test the complementarity of the carbon factor, we can compare the performance of traditional models like the CAPM, Carhart 4 factor, and Fama-French 3 factor models while incorporating the BMG factor. This is equivalent to evaluating the following CAPM+BMG model (Roncailli et al., 2020):

$$r_{i,t} - r_f = \alpha_i + \beta_i^{MKt}\left(R_t^{MKt} - r_f\right) + \beta_i^{BMG}R_t^{BMG} + \epsilon_{i,t} \tag{17}$$

and the following Carhart+BMG model:

$$r_{i,t} - r_f = \alpha_i + \beta_i^{MKt}\left(R_t^{MKt} - r_f\right) + \beta_i^{SMB}R_t^{SMB} + \beta_i^{HML}R_t^{HML} + \beta_i^{WML}R_t^{WML} + \beta_i^{BMG}R_t^{BMG} + \epsilon_{i,t} \tag{18}$$

where WML denotes the momentum factor.
If the BMG factor increases the explanatory power of these traditional models, it would suggest the BMG factor is capturing extra information. For this we can we perform the GRS test by (Gibbons et al., 1989) which tests whether the intercepts $\hat{\alpha}$ are indistinguishable from zero in the regression of the test assets' excess returns on the model's factor returns i.e.

- $\mathbf{H}_0$: $\alpha_i = 0 \ \forall i$
  $\mathbf{H}_\text{a}$:$\alpha_i \neq 0$

Under the null assumption, the GRS test assumes that the model is correct i.e. it effectively captures the cross-sectional returns of the underlying assets. However, as no model is perfect, we rather use the GRS test statistic and the level of α to compare the statistical power of models. Explicitely, models that better fit the data will have lower GRS statistics and smaller absolute level of $|\alpha|$. The GRS test can be done using the **GRS.test** package in R. Furthermore, we follow the methodology of (Gorgen et al., 2019) and use 5x5 sorted portfolios as test assets. 5x5 sorted portfolios follow the same methodology mentioned above, but we sort portfolios into 5 size groups and 5 groups sorted on an extra variable such as BE/ME, which yields 25 portfolios that we use as tests assets.

A complementary test to the GRS statistic is to use individual asset returns as test assets and calculate the mean adjusted R-squared. The adjusted R-squared measures how well a model fits the data, while accounting for unnecessary predictors. So if the BMG carbon factor is significant, models including it should have higher adjusted R-squared values on average (Brooks, 2014).

4.3.2 Fama-Macbeth procedure

The second condition for evaluating the consistency of the BMG factor is to test whether this portfolio is associated with a risk premium (excess return). The intuition stemming from the asset pricing theory, is that riskier assets should have higher expected returns over time compared to safer assets. This is because, investors should to be rewarded with higher potential returns for holding riskier assets in their portfolios instead of safer ones. If this theory hold for the carbon risk, we should observe a significant positive risk premium associated with the BMG factor.

To test this hypothesis we use to the 2 step regressions proposed by (Fama-Macbeth, 1973).

In a first step we estimate the beta exposures $\hat{\beta}_i^k$ using the same regression for (6) for the whole sample. Although in their original paper (Fama and Macbeth 1973) use a rolling window of 5 years to estimate the time t risk premium, we may use the whole sample under the assumption that the risk premium is constant over time.

In the second step regression, at each time t, we run cross-sectional regressions of the asset returns over the estimated betas obtained from the first step regressions. This is equivalent to running the following regression (Brooks, 2014):

$$\begin{bmatrix} r_{1,t} \\ \vdots \\ r_{n,t} \end{bmatrix} - r_f = \lambda_0 + \lambda_t^{MKt} \begin{bmatrix} \hat{\beta}_1^{Mkt} \\ \vdots \\ \hat{\beta}_n^{Mkt} \end{bmatrix} + \lambda_t^{SMB} \begin{bmatrix} \hat{\beta}_1^{SMB} \\ \vdots \\ \hat{\beta}_n^{SMB} \end{bmatrix} + \lambda_t^{HML} \begin{bmatrix} \hat{\beta}_1^{HML} \\ \vdots \\ \hat{\beta}_n^{HML} \end{bmatrix} + \lambda_t^{BMG} \begin{bmatrix} \hat{\beta}_1^{BMG} \\ \vdots \\ \hat{\beta}_n^{BMG} \end{bmatrix} + \epsilon_i \quad (19)$$

Where λ_t^k is the risk premium associated a factor k at time t. This yield a time series $\hat{\lambda} = [\hat{\lambda}_1^k, ...\hat{\lambda}_T^k]$ for each factor k, the average risk premium is then taken as the final estimator of $\hat{\lambda}^k$ i.e.

$$\hat{\lambda}^k = \frac{1}{T} \sum_{t=1}^{T} \hat{\lambda}_t^k \quad (20)$$

We can then test for significancy of risk premium using T-tests from the **stats** package in R, which is equivalent to calculating the following statistic for each factor k:

$$\frac{\hat{\lambda}^k}{\hat{\sigma}(\lambda^k)\sqrt{T}} \quad (21)$$

where T represents the sample size.

If the estimated parameters are significantly different from zero, we conclude that the factor is associated with a risk premium. Using the asset pricing jargon, we say that this risk is priced in the market. However, the Fama-Macbeth is prone to bias due to two main problems. First, it is prone to an Errors-in-variables problem. Since the second regression is run over estimates of $\hat{\beta}_i^k$ which are not observed, this would lead to biased estimations of $\hat{\lambda}^k$ (Pukthuanthong et al., 2019). Secondly, the Fama–MacBeth estimator ignores the serial correlation of returns (Jagannathan et al., 2009). Monthly asset returns often exhibit little variation, implying there might be autocorrelation from one month to another. This autocorrelation violates the assumption of independent returns and can bias the standard errors from the Fama-MacBeth asymptotic estimations. To overcome this issue, we can proceed with 2 modifications. First, following (Pukthuanthong et al., 2019), we use sorted portfolios as tests assets in the 1st and 2nd step. Specifically, we use monthly returns of industry portfolios sorted on the GICS sector and the 25 (5x5) portfolios sorted on size and Book market equity (BE/ME). Additionally, we can follow the work of (Jagannathan, et al. 2009) and estimate Newey-West heteroskedasticity and autocorrelation-consistent (HAC) standard errors of $\hat{\beta}_i^k$. This can be done using the **sandwich** to estimate the Newey-West covariance matrix of the risk premiums, then retrieve the associated standard deviations. Note that, using Newey-West standard errors do not change the parameters values of $\hat{\beta}_i^k$. However, it can yield better estimates of $\hat{\sigma}(\lambda^k)$ that better assess the statistical significancy of the risk premiums.

Thus we have defined two metrics for assessing a company's ESG profile: the carbon footprint and the carbon beta risk. While the latter will be used as an extra variable in the MV optimization framework, the carbon risk can be used as an extra metric to evaluate the financial risk that stems from investing in brown companies. In what follows, we introduce the Mean variance optimization frameworks that will be used to incorporate ESG criteria into portfolio allocation.

5 ESG integration: extending the Mean variance framework

We introduce two extensions of the general mean variance optimization framework that can be used to incorporate ESG related data into the asset allocation problem. The first one, consists of adding an ESG utility terms in the mean variance utility function and deriving the ESG efficient frontier following the methodology of (Pedersen et al., 2021). In a second approach, we define the MV optimization with a benchmark, which consists in building a portfolio which closely ressembles a benchmark portfolio while reducing its ESG score.

5.1 The ESG-efficient frontier

To illustrate the framework of (Pedersen et al., 2021), consider an agent initial wealth W and terminal wealth $\bar{\mathrm{W}} = (1 + \mathrm{r}_f + \pi^T\mu)\mathrm{W}$ where r_f is the risk free rate, $\boldsymbol{\mu}$ is the vector of expected **excess returns**, and $\boldsymbol{\pi} = [\pi_1, ..., \pi_n]'$ is the vector of **risky assets**. The objective of the investor is to maximize the utility from terminal wealth $\bar{W}$ by investing in a set of risky assets and the risk free rate.
The model uses the mean-variance utility function, which is tilted by the average ESG score of the portfolio:

$$u(\pi) = \mathbb{E}[\bar{\mathrm{W}}] - \frac{\bar{\gamma}}{2}\mathrm{Var}(\bar{\mathrm{W}}) + \bar{u}(\mathcal{S}^{port})\mathrm{W} \tag{22}$$

$$= \left(1 + r_f + \pi^T\mu - \frac{\bar{\gamma}}{2}\pi^T\Sigma\pi + \bar{u}(\frac{\pi^T\boldsymbol{\mathcal{S}}}{\pi^T\mathbf{1}})\right)\mathrm{W} \tag{23}$$

where $\boldsymbol{\mathcal{S}} = [\mathcal{S}_1, ..., \mathcal{S}_n]$ is the vector of ESG scores and $\mathcal{S}^{port}$ is the portfolio average ESG score (i.e. the weighted sum of ESG scores, scaled by the total position in risky assets). Similarly, we introduce a function $\bar{u}()$ which reflects the agent's preference for ESG investing. Here Σ is the covariance matrix, given by:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \dots & \sigma_{1n} \\ \vdots & \ddots & \vdots \\ \sigma_{n1} & \dots & \sigma_n^2 \end{bmatrix} \tag{24}$$

where $\sigma_{ij} = \mathrm{Cov}(r_i, r_j)$.
Optimizing the utility function yields the mean-variance-ESG optimized portfolio:

$$\boldsymbol{\pi}^* = \underset{\pi}{\mathrm{argmax}}\ \pi^T\mu - \frac{\bar{\gamma}}{2}\pi^T\Sigma\pi + \bar{u}(\frac{\pi^T\boldsymbol{\mathcal{S}}}{\pi^T\mathbf{1}}) \tag{25}$$

where the set of feasible portfolios must satisfy $\pi^T\mathbf{1} > 0$, that is all long-biased portfolios. That means portfolios must have a net positive position in risky assets, ruling out short-only portfolios. As explained by the authors, short-only portfolios do not have a straightforward interpretation for calculating the average ESG score. Thus, by limiting the optimization to long-biased portfolios, the average ESG score can be clearly defined based on the assets held in the portfolio.
In this setting, the investor aims to find the optimal portfolio that yields the best risk-return-ESG tradeoff. In contrast to traditional MV optimization, this problem cannot be solved using Karush–Kuhn–Tucker optimality conditions because of the additionnal ESG term. Although a set of different algorithms exist to fix this issue, including the Sequential Quadratic Programming (SQP) algorithms (Bert and Grauer, 1990).

The SQP algorithm provides an iterative optimization technique commonly applied to constrained mean-variance problems. It works by first solving a sequence of smaller subproblems with fixed parameters and binding constraints, then optimizing over all possible parameter values to approximate the original problem.

Let $\sigma = \sqrt{\pi^T \Sigma \pi}$ and $\mathcal{S} = \pi^T \mathcal{S}$, problem (25) can be decomposed as follows:

$$\max_{\bar{\mathcal{S}}} \left\{ \max_{\bar{\sigma}} \left\{ \max_{\pi} \ \pi^T \mu - \frac{\bar{\gamma}}{2} \pi^T \Sigma \pi + \bar{u}(\frac{\pi^T \mathcal{S}}{\pi^T \mathbf{1}}) \right\} \right\} \tag{26}$$

$$\text{subject to:}$$

$$\pi^T \Sigma \pi = \bar{\sigma}^2 \tag{27}$$

$$\frac{\pi^T \mathcal{S}}{\pi^T \mathbf{1}} = \bar{\mathcal{S}} \tag{28}$$

where $\bar{\sigma}^2$ and $\bar{\mathcal{S}}$ represent fixed values for the portfolio's volatility and average ESG levels. The idea here is to first choose the optimal portfolio $\pi^*(\bar{\sigma}, \bar{\mathcal{S}})$ that maximizes the objective function as a function of $\bar{\sigma}$ and $\bar{\mathcal{S}}$, then in a second place choose the optimal level of risk $\bar{\sigma}^*(\bar{\mathcal{S}})$ as a function of $\bar{\mathcal{S}}$, and finally choose the optimal level of ESG performance $\mathcal{S}^*$.
The first sub-problem is then given by:

$$\boldsymbol{\pi}^*(\bar{\sigma}, \bar{\mathcal{S}}) = \arg\max \pi^T \mu \tag{29}$$

$$\text{subject to:}$$

$$\pi^T \Sigma \pi = \bar{\sigma}^2 \tag{30}$$

$$\frac{\pi^T \mathcal{S}}{\pi^T \mathbf{1}} = \bar{\mathcal{S}} \iff \pi^T \tilde{\mathcal{S}} = 0 \tag{31}$$

where $\tilde{\mathcal{S}} = \mathcal{S} - \mathbf{1}\bar{\mathcal{S}}$.
The corresponding Lagrangien can be formulated as:

$$\mathcal{L}(\pi, \lambda, \alpha) = \pi^T \mu - \frac{1}{2} \lambda (\pi^T \Sigma \pi - \bar{\sigma}^2) + \alpha \pi^T \tilde{\mathcal{S}} \tag{32}$$

where the first order condition is given by:

$$\frac{\partial \mathcal{L}(\pi, \lambda, \alpha)}{\partial \pi} = \mu - \lambda \Sigma \pi + \alpha \tilde{\mathcal{S}} = 0 \tag{33}$$

rearranging (33), yields the optimal portfolio weights as a function of $\bar{\sigma}$ and $\bar{\mathcal{S}}$:

$$\pi^*(\bar{\sigma}, \bar{\mathcal{S}}) = \frac{1}{\lambda} \Sigma^{-1} \left(\mu + \alpha \tilde{\mathcal{S}} \right) \tag{34}$$

we refer to the authors' paper for the derivation of the optimal Lagrangien parameters which are given by:

$$\lambda = \frac{1}{\bar{\sigma}} \sqrt{C_{\mu\mu} - \frac{(C_{\mathcal{S}\mu} - C_{1\mu}\bar{\mathcal{S}})^2}{C_{\mathcal{S}\mathcal{S}} - 2C_{1\mathcal{S}}\bar{\mathcal{S}} + C_{11}(\bar{\mathcal{S}})^2}} \tag{35}$$

$$\alpha = \frac{C_{1\mu}\bar{\mathcal{S}} - C_{\mathcal{S}\mu}}{C_{\mathcal{S}\mathcal{S}} - 2C_{1\mathcal{S}}\bar{\mathcal{S}} + C_{11}\bar{\mathcal{S}}^2} \tag{36}$$

we use the notation $C_{xy} = x^T \Sigma^{-1} y$.
Using (34), the objective function can be rewritten as:

$$\max_{\bar{\mathcal{S}}} \left\{ \max_{\bar{\sigma}} \left\{ \pi^{*T}\mu - \frac{\bar{\gamma}}{2}\bar{\sigma}^2 + \bar{u}(\bar{\mathcal{S}}) \right\} \right\} \tag{37}$$

$$\text{subject to:} \quad \pi^* = \pi^*(\bar{\sigma}, \bar{\mathcal{S}}) \tag{38}$$

$$\iff \max_{\bar{\mathcal{S}}} \left\{ \max_{\bar{\sigma}} \left\{ \frac{\pi^*(\bar{\sigma}, \bar{\mathcal{S}})^T \mu}{\bar{\sigma}} \bar{\sigma} - \frac{\bar{\gamma}}{2}\bar{\sigma}^2 + \bar{u}(\bar{\mathcal{S}}) \right\} \right\} \tag{39}$$

$$\iff \max_{\bar{\mathcal{S}}} \left\{ \max_{\bar{\sigma}} \left\{ \mathrm{SR}(\bar{\mathcal{S}}|\pi^*, \Sigma, \mathcal{S})\bar{\sigma} - \frac{\bar{\gamma}}{2}\bar{\sigma}^2 + \bar{u}(\bar{\mathcal{S}}) \right\} \right\} \tag{40}$$

where $\mathrm{SR}(\bar{\mathcal{S}}|\pi^*, \Sigma, \mathcal{S})$ is the Sharpe ratio of the optimal portfolio $\pi^*(\bar{\sigma}, \bar{\mathcal{S}})$ i.e. the maximum Sharpe ratio that can be achieved with a score of $\bar{\mathcal{S}}$.
Multiplying the FOC condition (33) by π and rewriting yields:

$$\begin{aligned} 0 &= \mu - \lambda\Sigma\pi + \alpha\tilde{\mathcal{S}} \\ &= \pi^T\mu - \lambda\pi^T\Sigma\pi + \alpha\pi^T\tilde{\mathcal{S}} \\ &= \pi^T\mu - \lambda\bar{\sigma}^2 \\ \iff \lambda &= \frac{\pi^T\mu}{\bar{\sigma}}\frac{1}{\bar{\sigma}} \\ \lambda &= \frac{\mathrm{SR}(\pi, \bar{\sigma}, \bar{\mathcal{S}})}{\bar{\sigma}} \end{aligned} \tag{41}$$

Thus the Sharpe ratio of the optimal portfolio $\pi^*(\bar{\sigma}, \bar{\mathcal{S}})$ is given by:

$$\mathrm{SR}(\bar{\mathcal{S}}|\mu, \Sigma, \mathcal{S}) = \sqrt{C_{\mu\mu} - \frac{(C_{1\mu}\bar{\mathcal{S}} - C_{\mathcal{S}\mu})^2}{C_{\mathcal{S}\mathcal{S}} - 2C_{1\mathcal{S}}\bar{\mathcal{S}} + C_{11}\bar{\mathcal{S}}^2}} = \mathrm{SR}(\bar{\mathcal{S}}) \tag{42}$$

Having defined the optimal weights $\pi^*(\bar{\sigma}, \bar{\mathcal{S}})$, as a function of $\bar{\sigma}$ and $\bar{\mathcal{S}}$, the 2nd problem involves finding the optimal level of risk $\bar{\sigma}^*(\bar{\mathcal{S}})$ as a function of $\bar{\mathcal{S}}$:

$$\bar{\sigma}^*(\bar{\mathcal{S}}) = \underset{\bar{\sigma}}{\operatorname{argmax}} \left\{ \max_{\pi} \pi^T\mu - \frac{\bar{\gamma}}{2}\bar{\sigma}^2 + \bar{u}(\bar{\mathcal{S}}) \right\} \tag{43}$$

$$= \underset{\bar{\sigma}}{\operatorname{argmax}} \left\{ \mathrm{SR}(\bar{\mathcal{S}})\bar{\sigma} - \frac{\bar{\gamma}}{2}\bar{\sigma}^2 + \bar{u}(\bar{\mathcal{S}}) \right\} \tag{44}$$

The FOC condition yields $0 = \mathrm{SR}(\bar{\mathcal{S}}) - \bar{\gamma}\bar{\sigma}$ and:

$$\bar{\sigma}^*(\bar{\mathcal{S}}) = \frac{1}{\bar{\gamma}}\mathrm{SR}(\bar{\mathcal{S}}) \tag{45}$$

Using this result, we can further simplify the objective function to:

$$f(\pi^*(\bar{\sigma}, \bar{\mathcal{S}}), \mu, \Sigma, \mathcal{S}) = \mathrm{SR}^2(\bar{\mathcal{S}}) + 2\bar{\gamma}\bar{u}(\bar{\mathcal{S}}) \tag{46}$$

where $\bar{\gamma}$ now represents the ESG aversity of the agent. A higher ESG averse agent will put more weight on his ESG preferences and will prefer to sacrifice a part of his financial returns for a better ESG performance portfolio. The last step consists in finding the optimal $\mathcal{S}^*$ is found by solving the $\mathcal{S}$-problem given by:

$$\mathcal{S}^* = \underset{\bar{\mathcal{S}}}{\operatorname{argmax}} \left\{ \mathrm{SR}^2(\bar{\mathcal{S}}) + 2\bar{\gamma}\bar{u}(\bar{\mathcal{S}}) \right\} \tag{47}$$

and the optimal portfolio is given by:

$$\pi^* = \pi^*(\sigma^*, \mathcal{S}^*) \tag{48}$$

where $\sigma^* = \frac{1}{\gamma}\mathrm{SR}(\mathcal{S}^*)$.
This is the result concluded by (Pedersen et al., 2021). We can see that the problem, which initially was the optimization over 3 parameters μ, σ and $\mathcal{S}$ is reduced to optimizing over 1 parameter, $\mathcal{S}$.

In summary, the investor now chooses the optimal ESG score $\mathcal{S}^*$ that maximizes the function (46), which is then used to choose an optimal risk level $\sigma^* = \frac{1}{\gamma}\mathrm{SR}(\mathcal{S}^*)$. Finally, these values are used to compute the optimal Lagrange multipliers (35) and (36) which we substitute in the expression for the optimal portfolio weights displayed in equation (34).
More importantly, this result show that we can separate the ESG portfolio decision into two parts. One that depends only on the securities $\mathrm{SR}(\mathcal{S})$ and another that depends only on investor preferences $\bar{u}(\mathcal{S})$. Specifically, the ESG-Sharpe ratio frontier shows the maximum Sharpe ratio attainable for each level of ESG, independent of any investor preferences. The investor then chooses where to place themselves on this frontier based on their preferences over ESG and risk.

5.2 Portfolio optimization in presence of a benchmark:

The ESG-efficient frontier framework developed by (Pedersen et al., 2021) provides a method for selecting assets that offer the best risk-adjusted returns for a given ESG score. However, some investors may take simpler approaches to sustainable investing that are easier to implement. As such, asset managers may track progress on decarbonization using benchmarks rather than picking individual assets based on ESG-risk-return tradeoffs. The general framework for mean-variance optimization with a benchmark can be applied to incorporate ESG criteria in this way. In what follows, we define the general framework for the MV optimization with a benchmark and how it can be applied to incorporate an ESG criteria.

We define the benchmark portfolio $\boldsymbol{b} = [b_1, b_2, ..., b_n]'$ where b_i is the weight of asset i in the benchmark. Following (Roncailli, 2013), we introduce the notion of the tracking e for portfolio π which is the difference between the active return of the portfolio r^{port} and the return of the benchmark $r^{\text{benchmark}}$:

$$\begin{aligned} e &= r^{\text{port}} - r^{\text{benchmark}} \\ &= \pi^T r - b^T r \\ &= (\pi - b)^T r \end{aligned} \tag{49}$$

The expected excess return is given by:

$$\mu(\pi|b) = \mathbb{E}[e] \tag{50}$$
$$= (\pi - b)^T \mu \tag{51}$$

The tracking error volatility is defined as:

$$\sigma(\pi|b)^2 = \sigma(e)^2 \tag{52}$$
$$= (\pi - b)^T \Sigma (\pi - b) \tag{53}$$

where the $\sigma(\pi|b)$ measures how closely the portfolio replicates the benchmark. Thus minimizing the TE variance produces a portfolio that closely replicates the benchmark composition.

The objective of the investor is to maximize the expected excess return with a constraint on the tracking error or equivalently, to minimize the tracking error while constraining the expected excess returns. For the latter, this can be represented by the following optimization problem:

$$\min_{\boldsymbol{\pi}} \ (\pi - b)^T \Sigma (\pi - b) \tag{54}$$

$$\text{subject to:}$$

$$(\pi - b)^T \mu \geq \mu^* \tag{55}$$

$$\pi \in \Omega \tag{56}$$

where, μ^* is the desired level of excess return and Ω represents the set of additional constraints we want to impose.

This setting can be modified to account for ESG constraints by replacing the expected excess return of the potfolio π with the excess ESG performance defined as (Roncailli, 2023):

$$\mathcal{S}(\pi|b) = \pi^T \mathcal{S} - b^T \mathcal{S} \tag{57}$$

$$= (\pi - b)^T \mathcal{S} \tag{58}$$

The optimization problem then becomes:

$$\min_{\boldsymbol{\pi}} \ (\pi - b)^T \Sigma (\pi - b) \tag{59}$$

$$\text{subject to}$$

$$(\pi - b)^T \mathcal{S} \geq \mathcal{S}^* \tag{60}$$

$$\pi \in \Omega \tag{61}$$

where $\mathcal{S}^*$ is the desired level of excess ESG performance we want to achieve. The goal is to construct a portfolio that closely resembles the benchmark asset allocation, while increasing the portfolio's ESG score by $\mathcal{S}^*$ compared to the benchmark.

Alternatively, we can minimize the following mean variance utility function (Roncailli, 2023):

$$\min_{\boldsymbol{\pi}} \ f(\pi|b) = \frac{1}{2}(\pi - b)^T \Sigma (\pi - b) - \gamma (\pi - b)^T \mathcal{S} \tag{62}$$

$$= \frac{1}{2}(\pi^T \Sigma \pi - 2\pi^T \Sigma b + b^T \Sigma b) - \gamma (\pi - b)^T \mathcal{S} \tag{63}$$

$$= \frac{1}{2}\pi^T \Sigma \pi - \pi^T (\gamma \mathcal{S} + \Sigma b) + \frac{1}{2} b \Sigma b + \gamma b^T \mu \tag{64}$$

$$= \frac{1}{2}\pi^T \Sigma \pi - \pi^T (\gamma \mathcal{S} + \Sigma b) + c \tag{65}$$

where, $c = \frac{1}{2} b \Sigma b + \gamma b^T \mu$ is a constant that does not depend on the portfolio π, thus that can be excluded from the minimization problem. Here γ can be defined as the investor's ESG aversion, it controls the tradeoff between the portfolio's tracking error and its ESG improvement relative to the benchmark.

In their paper, (Le Guenedal and Roncailli, 2022) use the carbon intensity $\mathcal{CI}$ (5) as a measure of $\mathcal{S}$ in the optimization framework (59). They propose a slightly different ESG constraint than incorporates a reduction rate of the $\mathcal{CI}$ instead of using an absolute desired level $\mathcal{CI}^*$. Their framework is described as follows:

$$\pi^*(\mathcal{R}) = \text{argmin} \ \frac{1}{2}(\pi - b)^T \Sigma (\pi - b) \tag{66}$$

$$\text{subject to} \begin{cases} \pi^T \mathbf{1} = 1 \\ 0 \leq \pi_i \leq 1 \\ \pi^T \mathcal{CI} \leq (1 - \mathcal{R}) b^T \mathcal{CI} \end{cases} \tag{67}$$

where $\mathcal{R}$ represent the target carbon reduction rate compared to the benchmark carbon footprint. Note that the ESG constraint is reversed here, as opposed to (60), this is because lower levels of $\mathcal{CI}$ represent a better ESG performance.
This optimization problem involves constructing a portfolio π that closely resembles the benchmark portfolio b while decreasing carbon footprint by rate $\mathcal{R}$. Additional constraints may be imposed to reflect real-world limitations, such as long-only positions ($0 \leq \pi_i \leq 1$) to mimic institutional investors who often do not short stocks (Molt and Parknoy, 2019).

5.3 Testing for the stationarity of asset returns

In order to implement the mean variance framework, we need to estimate the parameters μ, σ, Σ using sample averages. However, if the underlying processes contain unit roots, this can violate the stationarity assumption and result in biased estimators. To avoid issues, we first need to test the stationarity of the test assets returns before using it for estimation and optimization. Specifically, we apply a confirmatory data analysis by considering two different tests, namely the Augmented Dickey Fuller (ADF) test (Fuller, 1976; Dickey and Fuller, 1979) and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test (Kwiatkowski et al., 1992).

Augmented Dickey–Fuller (ADF) test:

Below we summarize the Augmented Dickey-Fuller hypothesis testing:

- $\mathbf{H}_0$: series contains a unit root.
- $\mathbf{H}_\text{a}$: series is stationnary.

In addition, the lag parameter p has to be determined when applying the test. Though various selection methods exist, they can be arbitrary. Instead, a straightforward rule of thumb based on data frequency can be applied (Enders, 2014). For monthly data, 12 lags are reasonable. For quarterly data, 4 lags are suitable. The idea is to use a lag length corresponding to the periodicity of the data.
This test can be done in R, by using the **aTSA** package and specifying the lag parameter $p = 13$ to generate results up to lag 12.

Kwiatkowski–Phillips–Schmidt–Shin (KPSS):

Another commonly used test is the KPSS test (Kwiatkowski et al., 1992), which tests whether the time series is stationary against the alternative of the presence of a unit root.
Note that, a key difference from augmented Dickey-Fuller (ADF) tests is that the KPSS test has stationarity as the null hypothesis rather than the alternative hypothesis. The null and alternative hypotheses of the KPSS test are:

- $\mathbf{H}_0$: series is stationary.
- $\mathbf{H}_\text{a}$: series contains a unit root.

The KPSS test can be conducted in R using the **tseries** package. Using stationarity and unit root tests together provides more robust results and is known as confirmatory data analysis (Enders, 2014). However, for robust conclusions, the tests should yield similar results. If the tests produce conflicting outcomes, it can lead to inconclusive findings regarding stationarity. When stationarity and unit root tests draw the same conclusions, it provides stronger evidence for assessing the stationarity of a time series.

6 Empirical results

Where r_t and P_t denote the return and price at time t. This section presents the empirical results of our analysis. First, we assess the stationarity of the test assets (SN, SB, SG, BG) using the methods described in Section 5.2. Next, we estimate the carbon exposures for these assets following the approach outlined in Section 4.2. This involves evaluating the explanatory power of the BMG carbon factor using individual assets, then applying the Fama-Macbeth procedure to test the factor's consistency. Finally, we implement the ESG-integrated MV optimization and the MV optimization with a benchmark frameworks from Section 4.3, using (SN, SB, SG, BG) as tests assets.

6.1 Data:

We use the **Refinitiv database** to retrieve financial and carbon related data of the 1899 securities listed on the **STOXX Europe total market index**, which includes 1899 stocks as of 2023 and represents the Western European region as a whole.[12] Our sample spans from June 2007 until July 2022. For that period we retrieve the following data from the Refinitiv data, indicators in () show the identification codes of the mentioned variables in Refinitiv:

- monthly closing prices (P) from June 2007 until July 2022
- monthly market cap values (MV) for every year from June 2007 until June 2022
- End-year carbon intensity scores (ENERO03V) defined as $\frac{\text{Total scope 1,2 emission}}{\text{Revenues}}$ from 2006 until 2020
- End-year price-book value share (WC09304) from 2006 until 2020
- Company GICS sector name (TR.GICSSector)

Note that we do not include Scope 3 emissions in our analysis due to the scarcity of this data in the Refinitiv database for a large fraction of the companies listed in the Stoxx index. We also extract the European proxy for the monthly risk free rate over the sample period, from the Eugene Fama Data library.[13]
Finally, we generate the returns data using the following formula:

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{68}$$

6.2 Validating the BMG factor

In this subsection, we show results for validating the carbon risk factor as explained in Section (4.3). We start by estimating the model (6) using individual assets as test assets. In a first place, we generate the results for the Variance Inflation Factor (VIF) to test for the multicolinearity of the factors. Results shown in Table [6] show low results under 5, thus we don't need to adjust for multicolinearity in the data. This can be further investigated by plotting the correlation matrix related to the time series of each factor. For this we may also plot the correlation with the momentum factor WML which will be used further in this analysis. Table [**??**] shows low correlation overall between the factors. Additionally we observe a negative correlation of the BMG factor with respect to the other factors, indicating that it may contain extra information that explain the cross section variation of the returns. Additionally, we observe a high correlation between the momentum factor WML and the market factor MKT, however this factor will note be included in the estimation of model (6). We also generate the results for

[12] https://qontigo.com/index/bkxp/
[13] Availble here: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

MKT	SMB	HML	BMG
1.141664	1.054136	1.087018	1.011566

Table 6: This table shows the Variance Inflation Factor (VIF) test statistics estimated on a set of 1899 stocks from the STOXX Europe total market. The test is based on a panel data regression of the asset excess return on the Fama French 3 factor models augmented by the carbon BMG factor.The sample period ranges from July 2007 until June 2022. The global factors MKT, SMB, HML and the carbon factor BMG are constructed using European stocks from the STOXX Europe TM index.

	MKT	SMB	HML	WML	BMG
MKT	1				
SMB	0.204	1			
HML	0.277	0.028	1		
WML	0.970	0.392	0.283	1	
BMG	−0.090	0.009	−0.103	−0.098	1

Table 7: This table shows the correlation matrix of the 4 factors size (SMB), value (HML), momentum (WML) and carbon (BMG) factors constructed using the STOXX Europe TM index. The sample period ranges from July 2007 until June 2022.

the results for the Durbin and Watson test to check for serial collinearity of the returns. Table [8] shows the results for up to lag 12. The D-W test statistic close to 2 and p-values < 0.05 for all lags suggest that there is no need to correct for collinearity. Although, the regression with individual assets proves no need to correct for multicollinearity nor serial correlation, we also provide the results for the regression when using the tests assets as suggested with the Fama-Macbeth two step procedure as discussed in Section (4.3.2). Results are shown in Appendix (17) and display similar results with low VIF values at low lags, however colinearity becomes significant at lags higher than 4.

Lag	Autocorrelation	D-W Statistic	p-value
1	0.013	1.974	< 0.01
2	0.015	1.971	< 0.01
3	0.010	1.980	< 0.01
4	0.009	1.983	< 0.01
5	0.008	1.984	< 0.01
6	0.009	1.982	< 0.01
7	0.010	1.979	< 0.01
8	0.011	1.977	< 0.01
9	0.005	1.989	0.014
10	0.008	1.984	< 0.01
11	0.005	1.991	0.018
12	0.007	1.986	< 0.01

Table 8: This table shows the Durbin Watson test statistic and associated p-values to test for serial collinearity in monthly returns. The test was generated using a panel data regression of asset monthly excess returns from the STOXX Europe TM index onto 4 factor returns. The sample period ranges from July 2007 until June 2022. The global factors MKT, SMB, HML and the carbon factor BMG are constructed using European stocks from the STOXX Europe TM index.

In what follows, we evaluate the complementarity of the BMG factor with respect to traditional factors. For this we generate the results for the tests mentioned in Section (4.3.1).

Complementarity of the BMG factor

We start by generating the GRS test using 5x5 sorted portfolios as test assets. Specifically we follow the methodology of (Gorgen et al., 2019) and generate 5x5 portfolios sorted on size and Book to market equity, and 5x5 portfolios sorted on size and momentum. Next, we generate the GRS statistic for the (6), (17) and the (**??**) models. Results in Table [9]. Results show that models with the BMG factor have similar Mean $|\alpha|$, however these have a lower GRS statistic and a higher Mean R^2, except for the CAPM+BMG model in Panel A. This leads to assume that the BMG factor can explain these assets better than common models.

Model	GRS statistic	p-value	Mean R^2	Mean $\lvert\alpha\rvert$
Panel A: Size/Value 5x5 portfolios				
CAPM	2.542	0.000	0.313	0.069
CAPM + BMG	2.562	0.000	0.345	0.069
FF3	3.468	0.000	0.374	0.080
FF3 + BMG	3.445	0.000	0.407	0.080
Carhart	3.980	0.000	0.393	0.084
Carhart + BMG	3.980	0.000	0.393	0.084
Panel B: Size/Momentum 5x5 portfolios				
CAPM	2.527	0.000	0.313	0.073
CAPM + BMG	2.527	0.000	0.343	0.073
FF3	3.442	0.000	0.372	0.084
FF3 + BMG	3.418	0.000	0.404	0.085
Carhart	3.925	0.000	0.391	0.088
Carhart + BMG	3.925	0.000	0.391	0.088

Table 9: This table shows the results of various asset pricing tests on global test assets. We include 25 global portfolios formed on Size/Value and Size/Momentum formed using the STOXX Europe TM index. We compare various models with and without the BMG factor, according to the GRS test. The sample period ranges from July 2007 until June 2022. The global factors MKT, SMB, HML and the carbon factor BMG are constructed using European stocks from the STOXX Europe TM index.

To provide further results on the complementarity of the BMG factor. We follow the methodology in Section (4.3.2) and estimate the models used in Table [9] on individual assets. We calculate difference in the adjusted-R^2 for models with and without the BMG factor and average over all the assets to have an overall evaluation of the factor. Furthermore, we generate partial F-tests for each asset in each model and calculate the proportion of assets for which the F-test generates significant results at the 1%, 5% and 10% significance level. Results are shown in Table [10].

	Average(Δadj $-$ R^2)	1%	5%	10%
FF3 - Carhart	0.69	8.54	18.01	25.56
CAPM - CAPM + BMG	0.51	7.50	14.56	20.85
FF3 - FF3 + BMG	0.52	8.10	15.33	22.71
Carhart - Carhart + BMG	0.58	7.99	14.89	22.99

Table 10: This table shows the difference in adjusted R-squared $\Delta\text{adj} - \text{R}^2$ for traditional factors augmented by the carbon BMG factor and their nested models. The average adjusted R-squared is calculated as the average difference over all assets. 1%, 5%, and 10% represent the proportion of assets for which the partial F-test on nested models, rejected the null hypothesis of the BMG factor being insignificant. Asset used for evaluation include the 1899 constituents of the STOXX Europe TM index. The sample period ranges from July 2007 until June 2022. The global factors MKT, SMB, HML and the carbon factor BMG are constructed using European stocks from the STOXX Europe TM index.

The average increase of the FF3-4F adjusted R^2 when adding the momentum factor is 0.69%. This increase is significant for 18.00% of the firms in the sample at the 5% significance level. In comparison, the BMG factor increases the adj. R2 by 0.52% and is significant for 15.33% of the assets. Additionally, the BMG factor increases the explanatory power of the 4F model by 0.58% and is significant for 14.89% of the assets at the 5% confidence level, showing that the BMG factor can potentially increase the explanatory power of traditional factor models.

Fama-Macbeth results

Finally, we proceed with the Fama-Macbeth two step procedure. We first estimate the Beta exposures $\hat{\beta}_i^k$ from model (6) with a test portfolio containing the 25 value weighted portfolio sorted on size and book to market equity, and 10 industry portfolios sorted on GICS. In a second step we use these estimates as independant variables and the tests assets returns in a panel data regression. Table [11] shows that overall, none of the estimated risk premia are significant. The carbon risk premium is -0.18%, but insignificant and therefore lacks the ability to explain the cross-section variation in excess returns. This aligns with the results of (Gorgen et al., 2019) which conclude that the BMG factor can indeed increase the statistical power of traditional factor models, however the carbon risk do not present a risk premium. This means that, based on the sample we used, we can conclude that investors do not seem to require additional compensation for their exposure to climate transition risk. In what follows,

Risk premium	Coefficient	p-value	Std
const	0.024	0.137	0.016
Mkt	-0.004	0.773	0.013
HML	0.012	0.432	0.015
SMB	0.019	0.379	0.021
BMG	-0.018	0.270	0.016

Table 11: This table show the risk premiums estimated using the Fama-Macbeth 2 step regressions. The estimation is based on a set of 35 portfolios. This include 25 portfolios sorted on size/Book market equity and 10 industry portfolios sorted based on the GICS sector. Assets used for the sorting include the 1899 constituents of the STOXX Europe TM index. p-value are obtainted from a two-sided T-test to test for the significancy of the risk premiums. The sample period ranges from July 2007 until June 2022. The global factors MKT, SMB, HML and the carbon factor BMG are constructed using European stocks from the STOXX Europe TM index.

we derive the results of the optimal portfolios first in the (Pedersen et al., 201) framework using the WACI metric (4) as a metric for ESG performance. We explain the intuition behind this methodology by deriving the optimal portfolios for two types of investors as postulated by the authors. In a second stage, we present the portfolio decarbonization framework, which stems from a simpler and more intuitive framework that can help investors choose the optimal portfolio in terms of ESG-risk performance.

6.3 ESG-efficiet frontier

In order to investigate the ESG-efficient frontier, we need to define an investment universe. To keep tractability of the problem, we restrict to using the constituents of the BMG factor i.e. the SB, SG, BB and BG portfolios mentioned in Section (4.3.1) as test assets. Additionally, for each sorting period t we allocate a carbon footprint level to the sorted portfolios using the WACI method i.e.

$$\text{WACI}_t(\pi_j) = \sum_{i=1}^{n_j} \pi_{i,t} \cdot \frac{\mathcal{CE}_{i,t}}{\text{Revenues}_{i,t}} \tag{69}$$

where $j \in \{\text{SB}, \text{SG}, \text{BB}, \text{BG}\}$.
In a first place, we test for the stationarity of the returns associated with these portfolios. As mentioned in Section (5.3) we generate the results for the KPSS test and 2 versions of the ADF test including, the model with a drift term and trend term and the model with no drift term and no trend term. Results are summarized in table [12] and table [13].

	statistic	**p-value**
SB	0.174	0.1
SG	0.188	0.1
BB	0.181	0.1
BG	0.224	0.1

Table 12: This table shows the KPSS stationarity test statistic and the associated p-value obtained from the time series of 4 portfolios sorted on size and carbon intensities. The original sorting generated in 6 portfolios sorted on size (market capitalization) and carbon intensity including the SB, SN, SG, BB, BN and BG portfolios. For this test SN and BN were excluded from the analysis. Assets used for the sorting include the 1899 constituents of the STOXX Europe TM index. The sample period ranges from July 2007 until June 2022.

	SB				**SG**				**BB**				**BG**			
	no drift no trend		with drift and trend		no drift no trend		with drift and trend		no drift no trend		with drift and trend		no drift no trend		with drift and trend	
lag	statistic	p-value	statistic	p-value	statistic	p-value	statistic	p-value	statistic	p-value	statistic	p-value	statistic	p-value	statistic	p-value
0	-11.43	**0.01**	-11.58	**0.01**	-11.07	**0.01**	-11.33	**0.01**	-11.24	**0.01**	-11.25	**0.01**	-10.75	**0.01**	-10.83	**0.01**
1	-9.41	**0.01**	-9.65	**0.01**	-9.09	**0.01**	-9.45	**0.01**	-8.89	**0.01**	-8.92	**0.01**	-9.61	**0.01**	-9.75	**0.01**
2	-6.66	**0.01**	-6.89	**0.01**	-6.35	**0.01**	-6.68	**0.01**	-6.83	**0.01**	-6.88	**0.01**	-6.45	**0.01**	-6.58	**0.01**
3	-5.21	**0.01**	-5.43	**0.01**	-5.18	**0.01**	-5.51	**0.01**	-5.82	**0.01**	-5.90	**0.01**	-4.84	**0.01**	-4.97	**0.01**
4	-4.89	**0.01**	-5.13	**0.01**	-4.55	**0.01**	-4.89	**0.01**	-5.49	**0.01**	-5.57	**0.01**	-4.58	**0.01**	-4.68	**0.01**
5	-4.95	**0.01**	-5.24	**0.01**	-4.57	**0.01**	-4.96	**0.01**	-5.41	**0.01**	-5.51	**0.01**	-5.06	**0.01**	-5.20	**0.01**
6	-4.49	**0.01**	-4.78	**0.01**	-4.18	**0.01**	-4.60	**0.01**	-5.14	**0.01**	-5.17	**0.01**	-4.81	**0.01**	-4.87	**0.01**
7	-4.08	**0.01**	-4.38	**0.01**	-3.75	**0.01**	-4.19	**0.01**	-4.49	**0.01**	-4.56	**0.01**	-4.54	**0.01**	-4.62	**0.01**
8	-3.91	**0.01**	-4.23	**0.01**	-3.39	**0.01**	-3.86	**0.02**	-4.41	**0.01**	-4.45	**0.01**	-3.99	**0.01**	-4.06	**0.01**
9	-3.69	**0.01**	-4.03	**0.01**	-3.16	**0.01**	-3.63	**0.03**	-4.30	**0.01**	-4.42	**0.01**	-3.63	**0.01**	-3.76	**0.02**
10	-3.62	**0.01**	-3.98	**0.01**	-2.95	**0.01**	-3.44	**0.05**	-4.02	**0.01**	-4.15	**0.01**	-3.70	**0.01**	-3.80	**0.02**
11	-3.22	**0.01**	-3.59	**0.04**	-3.14	**0.01**	-3.71	**0.02**	-4.08	**0.01**	-4.18	**0.01**	-3.71	**0.01**	-3.76	**0.02**
12	-3.42	**0.01**	-3.83	**0.02**	-3.05	**0.01**	-3.65	**0.03**	-4.36	**0.01**	-4.41	**0.01**	-3.94	**0.01**	-4.02	**0.01**

Table 13: This table shows the Augmented Dicket Fuller stationarity test statistic and the associated p-value obtained from the time series of 4 portfolios sorted on size and carbon intensities. Left side test include the test with no drift and no trend term and the right side include the test with a drift term and a trend term. The original sorting generated in 6 portfolios sorted on size (market capitalization) and carbon intensity including the SB, SN, SG, BB, BN and BG portfolios. For this test SN and BN were excluded from the analysis. Assets used for the sorting include the 1899 constituents of the STOXX Europe TM index. The sample period ranges from July 2007 until June 2022.

Looking at the p-values generated by ths KPSS and the two version of the ADF test, we can conclude that the time series generated by the returns of the 4 portfolios are stationary, thus we may use the sample moments as measures of μ, σ and Σ for our analysis.

We can then generate some descriptive the descriptive statistics of the assets to get some first insights. Results are shown in table [14]. All portfolios show positive average returns, ranging

	mean	std	median	min	max
SB	0.570%	5.963%	1.173%	−22.580%	25.230%
SG	0.595%	5.172%	0.842%	−20.760%	20.010%
BB	0.021%	4.784%	0.528%	−16.780%	15.540%
BG	0.009%	5.913%	0.614%	−23.140%	27.910%

Table 14: This table shows summary statistics for the monthly excess returns of the 4 portfolios formed on size and carbon intensities. The original sorting generated in 6 portfolios sorted on size (market capitalization) and carbon intensity including the SB, SN, SG, BB, BN and BG portfolios. Assets used for the sorting include the 1899 constituents of the STOXX Europe TM index. The sample period ranges from July 2007 until June 2022.

from 0.009% for big green firms (BG) to 0.595% for small green firms (SG). The carbon efficient big firms portfolio (BG) shows higher volatility (5.913%) than the carbon inefficient big firm portfolio (BB) which has volatility of 4.784%. This relationship is reversed for small firms, with the carbon efficient portfolio (SG) having lower volatility of 5.172% than the brown one (SB) at 5.963%. The small firm portfolios (SB, SG) have substantially, higher average returns than the big firm ones (BB, BG).
We can also generate the portfolios yearly carbon footprints of these portfolios. Figure displays (2) displays the annual carbon footprint of each portfolio sorted from 2007 to 2021 using the WACI measure displayed in (4). Units are in kilotonnes of CO2/$M of revenues generated. At first sight, we see that brown firms show substantially higher carbon footprints compared to green firms, with values ranging between 0.3ktCO2/$M and 0.7ktCO2/$M, while green firms show values close to 0. In addition, we see that big brown firms (BB) show higher carbon footprints compared to small brown firms (SB), however this trend seems to invert since 2018.

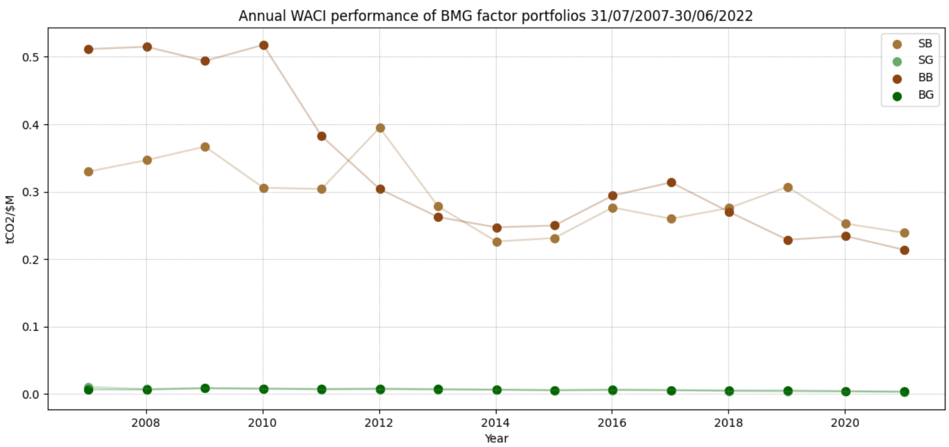


Figure 2: This figure shows the annual carbon footprint of of the 4 portfolios formed on size and carbon intensities. The original sorting generated in 6 portfolios sorted on size (market capitalization) and carbon intensity including the SB, SN, SG, BB, BN and BG portfolios. Assets used for the sorting include the 1899 constituents of the STOXX Europe TM index. Carbon footprint is allocated in each sorting date, using the WACI (4).The sample period ranges from July 2007 until June 2022.

To run the analysis of the ESG efficient frontier, we use the whole sample to have better

estimates of the μ. σ^2 and Σ. However, for the vector of carbon footprint $\mathcal{S}$ measured using the WACI measure, we use the latest values available i.e. the WACI values for the portfolio sorted in June 2021. We assume it makes more sense to use latest available carbon data than estimating the expected carbon footprint as the sample average. This comes from the fact that companies tend to reduce their carbon intensity over time, especially in the recent years. Thus using relatively old carbon data, provides reduced information about the future carbon intensity of a company.
In contrast, we can use the past carbon data in a statistical model to forecast the future carbon footprint using OLS estimation or the Kalman filter as suggested by (Le Guenedal et al., 2022). We investigate the relationship between ESG performance, measured by WACI, and the maximum attainable Sharpe ratio given by (42). Figure (3), plots the maximum attainable sharpe ratios as a function of $\bar{\mathcal{S}}$.

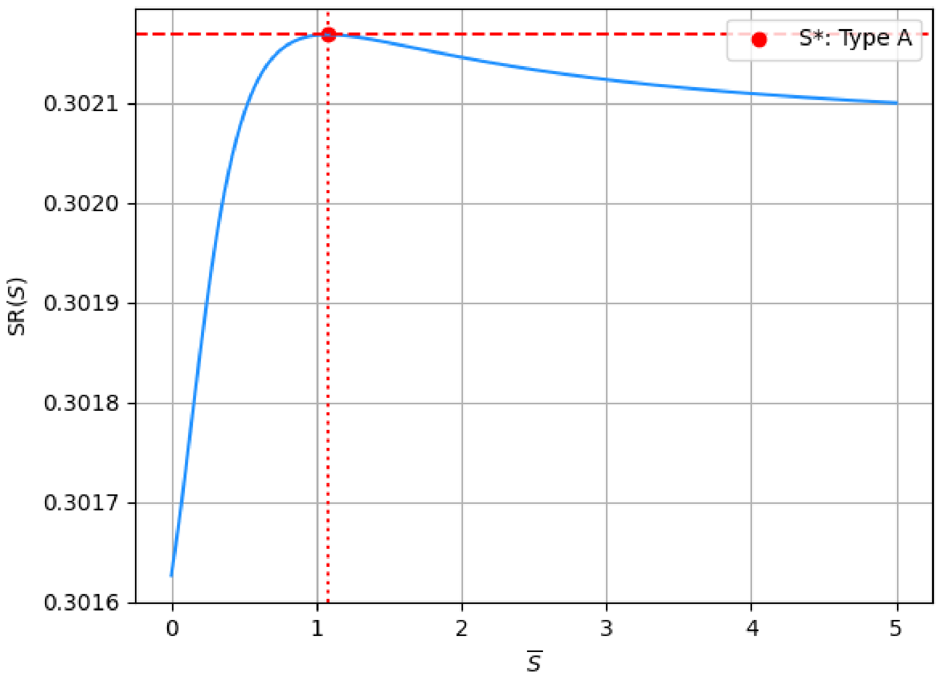


Figure 3: This figure shows the relationship between the portfolio carbon footprint ($\bar{\mathcal{S}}$) and the maximum attainable Sharpe ratio SR($\bar{\mathcal{S}}$). Carbon footprint is calculated using the WACI measure (4). Units are in kilotonnes of CO2/\$M revenues, based on carbon data provided by Refinitiv. $\mathcal{S}^*$ indicates the optimal ESG score for type A investors who have mean-variance preferences but use ESG data to update their views on risk and expected return. The curve illustrates the efficient frontier that trades off carbon intensity of the portfolio (ESG score) versus the maximum risk-adjusted return (Sharpe ratio). Assets used for the sorting include the 1899 constituents of the STOXX Europe TM index. The sample period ranges from July 2007 until June 2022.

We see that, as $\bar{\mathcal{S}}$ increases, the maximum attainable Sharpe ratio also increases until attaining a peak which corresponds to the maximum attainable Sharpe ratio. It is worth noting here that a higher carbon intensity corresponds to a worse ESG performance, since companies are less carbon efficient. Thus, ESG-motivated investors prefer portfolios on the left side of the curve with lower carbon intensity and better ESG scores. However, moving too far left on the curve eventually reduces the Sharpe ratio, which translates the tradeoff between ESG performance and risk-adjusted returns as noted by (Pedersen et al., 2021).

We expect that investors will not choose a point on the right side of the curve, as these portfolios will have a worse ESG score with no extra return benefit. Thus, they can aim for the same sharpe ratio while decreasing their carbon footprint.
The peak Sharpe ratio is achieved at a carbon intensity of approximately 1.08 ktCO2/$M. This represents the $\mathcal{S}$ score associated with the portfolio that has the maximum Sharpe ratio and coincides with the optimal portfolio for type A investors who have no ESG preference, but use the ESG scores to update their views on the risk-return trade-off.
Building on the framework from (Pedersen et al., 2021), we can examine the optimal strategy for ESG-motivated type M investors who incorporate ESG preferences into their investment decisions. To illustrate this visually, we plot the function (4) for a range of $\bar{\mathcal{S}}$. Specifically, two different ESG utility functions are used: a concave quadratic function and a convex logarithmic function. These utility specifications represent different investor preferences over ESG, where the concave functions imply the investor values initial ESG improvements more (decreasing marginal utility), while convex functions place higher weight on extra ESG gains at higher levels (increasing marginal utility). Note that, we negate these utility functions, to illustrate the fact that lower levels of $\bar{\mathcal{S}}$ generate better ESG utility. Results are shown in figure (4).

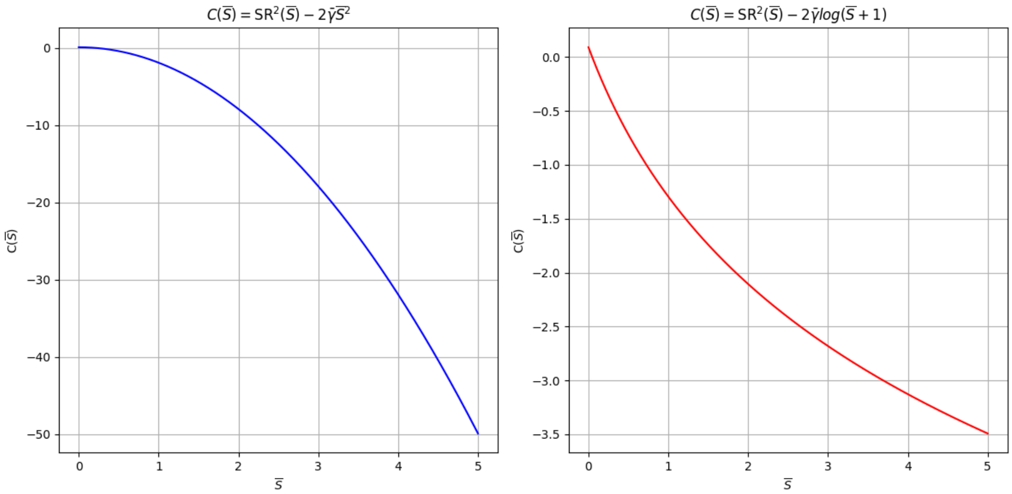


Figure 4: This figure shows the relationship between the carbon footprint ($\bar{\mathcal{S}}$) and function $C(\bar{\mathcal{S}})$ (46) which investor of type M optimizes to choose the optimal ESG score $\mathcal{S}^*$. Carbon footprint is calculated using the WACI measure (4). Units are in kilotonnes of CO2/$M revenues, based on carbon data provided by Refinitiv. Assets used for the sorting include the 1899 constituents of the STOXX Europe TM index. The sample period ranges from July 2007 until June 2022.

We see that for this example the optimal level for $\mathcal{S}$ is equal to 0 when using the concave function, which reduces considerably the carbon intensity of the portfolio compared to type-A investor. However, as type-M investor is constrained by its $\mathcal{S}$ scores, it naturally yields lower levels of Sharpe ratio.
In what follow, we investigate the weights corresponding to the optimal portfolios for type A and type M investors. We calculate the optimal Lagrangien parameters for $\mathcal{S} = 0$KtCO2/$M and $\mathcal{S} = 1.08$KtCO2/$M, and substitute in the formula for the optimal weights (34). Results are shown in Table (15).

	π^*_{SB}	π^*_{SG}	π^*_{BB}	π^*_{BG}	μ_π	$\mathcal{S}(\pi)$	SR(π)	r_f
Type A	7.589	8.468	−7.49	−8.376	9.13 %	1.08	0.302	0.808
Type M	6.917	9.061	−7.718	−8.085	9.09 %	0.0	0.301	0.825

Table 15: This table summarizes the results generated for the optimal weights using the framework of (Pedersen et al., 2021). Results were generated for 2 values of $\bar{\mathcal{S}}$, 0 (investor M) and 1.08 (investor A). Financial and carbon related data are the same used for the figure (**??**). π_i^* represents the optimal portfolio weight associated with each asset. μ_π and SR(π) represent the excess return and the Sharpe ratio associated with each strategy, while r_f represent the weight invested in the risk free asset and is equal to 1-$\sum \pi_i$.The sample period ranges from July 2007 until June 2022.

Upon initial examination, the two portfolios seem to be concentrated in a small number of highly leveraged positions. This could be due to inaccurate estimates of expected returns or covariance, or a significant difference in scale between the two metrics.
However, when we look at the investment strategies, we see that the portfolios have similar approaches. Both investors A and M invest heavily in small companies (SG and SB), which have a favorable risk-return profile compared to bigger companies (BB and BG). They also take large shortselling positions in BB and BG, with exposures of -7.49 and -8.376 for investor A and -7.718 and -8.085 for M, respectively. This is likely due to the low returns of these assets and high volatility of the BG portfolios.
Additionally, we can further investigate how investor M's investment strategy changes as he shifts to lower values of $\bar{\mathcal{S}}$ (lower ESG score). For this we recalculate the optimal portfolios for a range of values of $\bar{\mathcal{S}}$ ranging from 0 to 1.08, while keeping all other factors constant. Results are shown in fig (5).

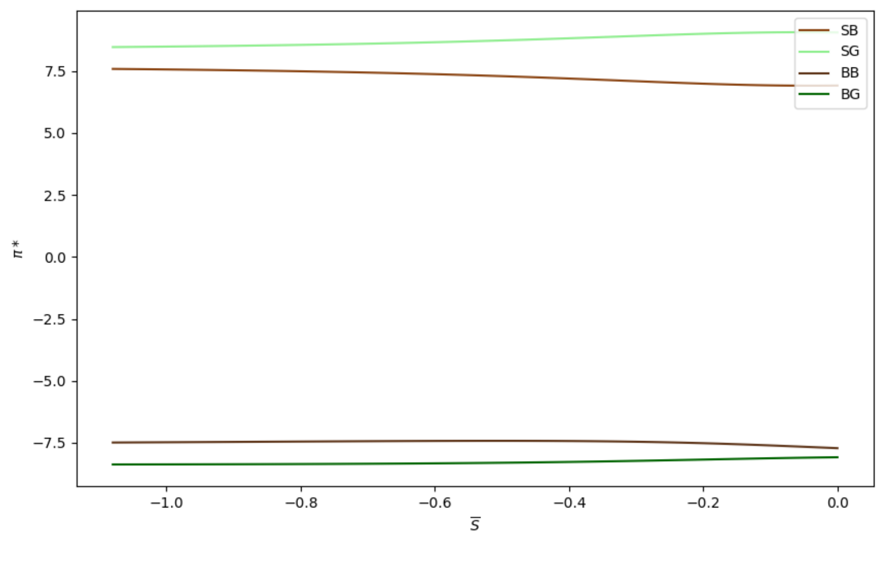


Figure 5: This figure shows how the optimal portfolio weights change for different values of the carbon footprint $\bar{\mathcal{S}}$. Higher values of $\bar{\mathcal{S}}$ (right side) indicate better ESG performance and are plotted on the x-axis. Here, we negate the carbon intensity metric to reflect higher values means better ESG performance. The optimal weights are generated using the parameters estimated in the portfolio analysis section, with $\bar{\mathcal{S}}$ varied across its range. The results demonstrate how the asset allocation shifts as the investor targets portfolios with higher ESG scores, represented by larger values on the x-axis.

As investor M imposes a stricter ESG score (moving from the left to the right side of the curve), they will prefer to increase their exposure to SG and shortselling position in BB assets.

However, we also see that even though the exposure to SB companies decreases, it remains a significant part of the portfolio. This can be explained by the fact that SB companies offer a risk-return tradeoff, despite having a higher carbon intensity score than green firms. In other words, investor M is willing to accept a higher carbon footprint from SB companies in order to achieve a higher return on their investment.
Similarly, we can evaluate the portfolios' exposures to the carbon risk mentioned in Section (4.1) in order to get additional insights about the transition risks associated with the optimized portfolios. For this we estimate asset-level carbon exposures as used in Section (6.2). Specifically, we estimate the model with individual assets and perform a panel data regression as mentioned above. This yield a set of beta estimate $\hat{\beta}_i^{\text{BMG}}$ for each asset. Next we allocate the portfolio level carbon exposures using the formula (70):

$$\mathcal{S}(\pi) = \sum_{i=1}^{n} \pi_i \cdot \hat{\beta}_i^{\text{BMG}} \tag{70}$$

This is done for for the assets included in the last sorting of (SB, SG, BB, BG).
Figure (6) displays the results. We can see that assets BG and SG yields negative carbon betas exposures, meaning that these companies should potentially benefit the most to shifting to a greener economy. Additionally, the carbon risk exposures show some differences compared to the raw WACI metric. Asset SB has a negative carbon beta exposure but asset BB has a positive exposure. This suggests BB will be more negatively affected by the low-carbon transition compared to SB, despite comparable carbon footprints over the sample as shown in Figure (2). So while SB and BB look similar based on past emissions, their carbon risk exposures indicate BB faces much higher climate transition risk going forward. This suggests that, carbon beta provides an additional forward-looking measure compared to carbon intensity alone. Overall, investor M's optimized portfolio has a more negative carbon risk exposure than investor A's. While investor A may have higher risk-adjusted returns, investor M's holdings seem less exposed to carbon transition risks, suggesting that investor M's portfolio could exhibit lower variability from the low-carbon shift.

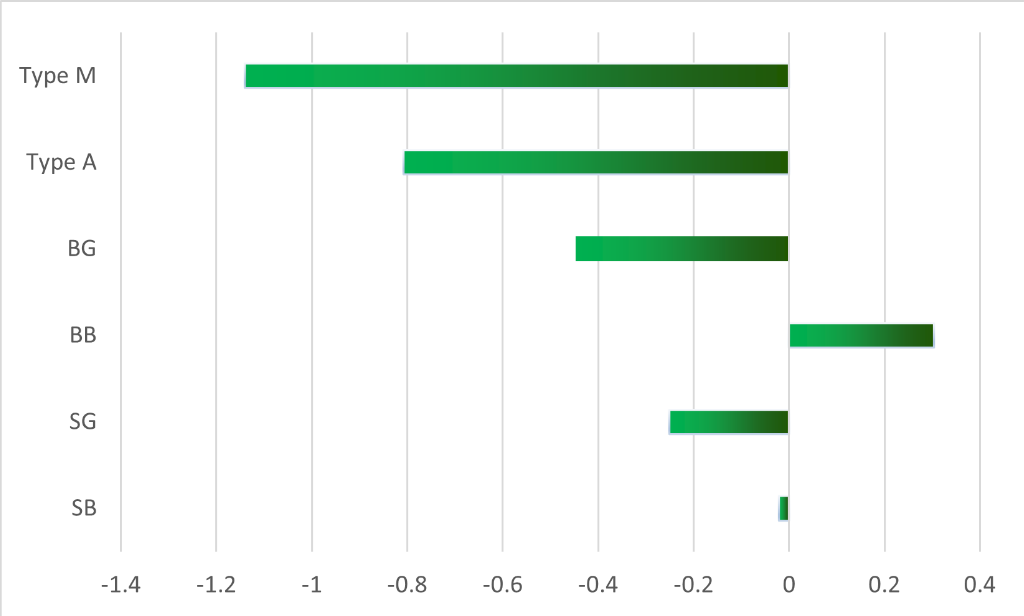


Figure 6: This figure shows the portfolio carbon exposures for the 4 portfolios sorted on size and carbon footprint (SB, SG, BB, BG), and for the optimal portfolio assets for type A and type M investors displayed in Table [15]. the x-axis shows the estimated carbon exposure $\hat{\beta}^{BMG}$ for each portfolio. We estimate asset level carbon exposures by a panel data regression of asset excess returns on the MKT, SMB, HML and carbon BMG factor as done in Section (6.2) and allocate portfolio level carbon exposures using formula (70). The estimation is done using the whole sample from July 2007 until June 2022.

Instability of the ESG-efficient frontier

The ESG-efficient frontier proposed by (Pedersen et al., 2021) can also yield unstable results depending on the data used. Specifically, certain optimized portfolios do not fulfill the condition long-biased portfolios i.e. $\pi^T \mathbf{1} > 0$. To illustrate this, we generate the optimal portfolios for both investors using an extended window approach. For each year t, starting from 2010 we estimate the μ, σ and Σ using all the prior observations, while using the latest available values of $\mathcal{S}$ prior to t. We then run the optimization problem and generate the results for $\sum_{i=1}^{n} \pi_i$ in Figure (7).

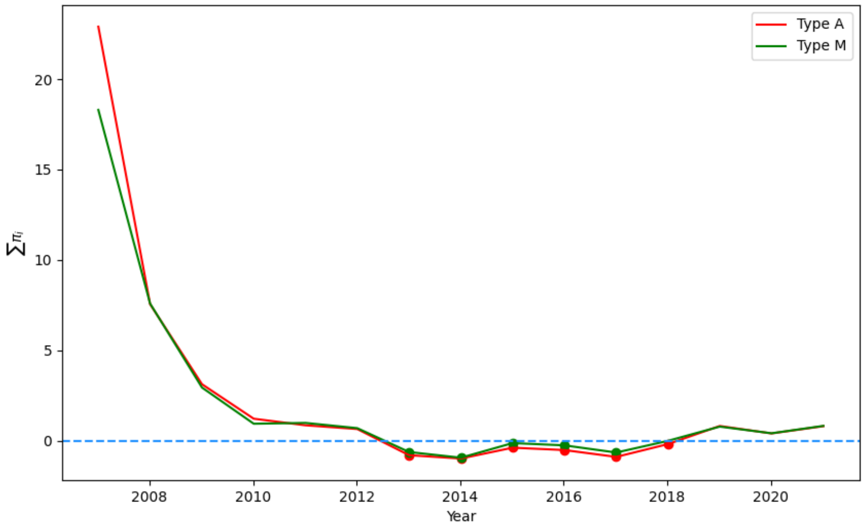
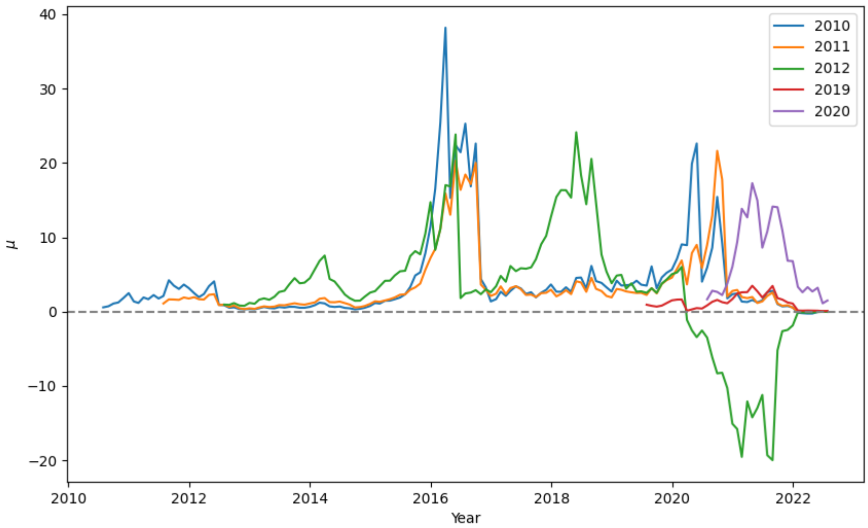


Figure 7: This figures shows the sum of the weights in risky assets for type A and type M investors using an extended sample under the (Pedersen et al., 2021) framework. Starting from 2007, each year we extend the sample, allocate a carbon footprint score using (4) and calculate the optimal portfolio weights for both investors using the methodology in Section (5.1). We then extract the sum of the weights $\sum_i^n \pi_i$ and plot them. Dots show years for which the optimal portfolios do not satisfy of long-biased portfolios.

Results show that, optimized portfolios between 2013 and 2018, do not satisfy the condition of long-biased portfolios for both type A and type M investors. Similarly portfolios constructed prior to 2010 yield very uncommon results with very high leveraged positions reaching up to 2000% leverage for 2008. Thus we cannot use the (Pedersen et al., 2021) framework to draw conclusion on the optimal weights for these portfolios.

Portfolio evaluation

For the remaining sample years (2010, 2011, 2012, 2019 and 2020), we generate the optimal portfolios for Investor A and Investor M and track the cumulative returns of these optimized portfolios over time. Figure (8) summarizes the results. We find that most portfolios yielded positive excess returns over the years, with portfolios constructed in 2010 and 2011 reaching approximately 3500% cumulative excess returns by 2016. While seemingly high, these excess returns are justified by the highly leveraged positions taken on by Investors A and M, which also translate into higher risk. As seen for 2012 portfolios, both investors endured severe losses from 2016 to 2020, with cumulative returns dropping to around -2000% for Investor A in 2021 and -1000% for Investor M in 2019. Overall, the optimized portfolios show strong potential upside but also significant downside risk inherent in leveraged investment strategies.



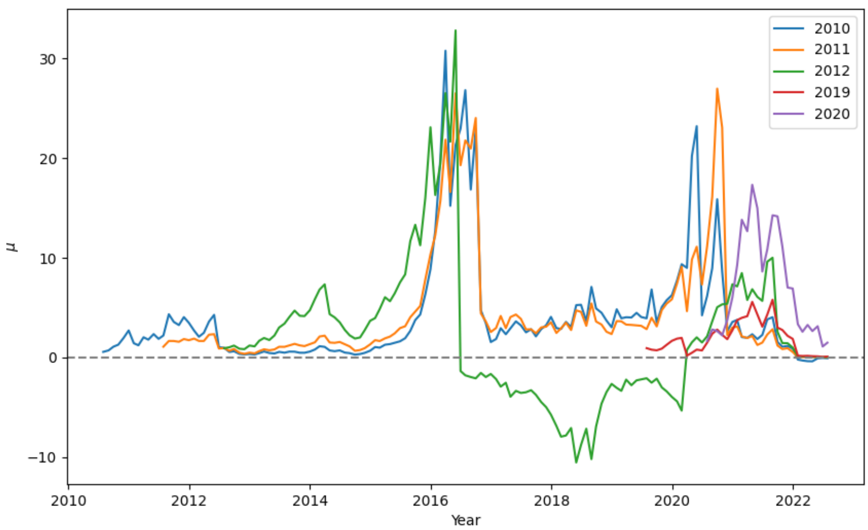


Figure 8: This figures shows the cumulated returns μ of the optimal portfolios for type A and type M investors. For the years (2010, 2011, 2012, 2019, 2020), we use all the prior data and perform portfolio optimization following the framework of (Pedersen et al., 2021). We then plot the cumulated monthly excess returns over time. Top: shows the cumulated monthly excess returns for type A investor. Bottom: shows the cumulated monthly excess returns.

Similarly we can generate the difference in the accumulated returns for investors A and M to track the performance of type A investor relative to the excess returns of type M investor. For this, we compute the difference in accumulated returns of investor's A portfolio relative to investor's M portfolio return and plot the results in Figure (9).

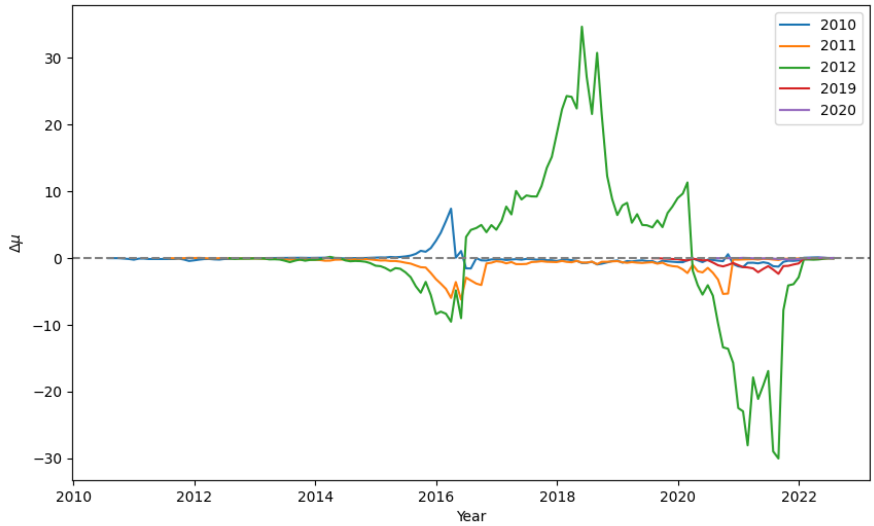


Figure 9: This figures shows the difference in cumulated returns $\Delta\mu$ between the optimal portfolios for type A and type M investors. For the years (2010, 2011, 2012, 2019, 2020), we use all the prior data and perform portfolio optimization following the framework of (Pedersen et al., 2021). We then calculate the cumulated monthly excess returns of both investors over time. For each year we calculate the difference in cumulated returns for each portfolio constructed in year t.

Figure (9) shows mitigated results for both investors. The portfolio constructed in 2012 exhibits the highest volatility in excess return differences. From 2016 to 2020, Investor A had substantially higher cumulative returns for the 2012 portfolio compared to investor M. However, after 2020, we see most of Investor M's portfolios outperforming those of Investor A. This suggests that post 2020, investing in greener portfolios generated excess returns compared to strategies solely focused on risk-adjusted return maximization. In fact, the latter strategies produced more volatile portfolios versus more sustainable ones. Overall, this suggests that incorporating ESG factors led to more stable excess returns during the recent period, indicating potential benefits of sustainable investing practices.

6.4 Portfolio decarbonization

As mentioned in Section (4.2.2), investors may use benchmarks to track progress on decarbonization objectives rather than selecting individual assets based on ESG-risk-return tradeoffs. To illustrate this, we apply the optimization problem (66) to the optimized portfolio obtained for Type A investor. We do not impose total allocation constraints nor long only constraints since the portfolio A does not satisfy those conditions.
We generate results of the optimization problem (66) for different levels of the decarbonization rate $\mathcal{R}$ and plot the results in Table (16).
Results are similar to those obtained with the ESG efficient frontier. As we increase the decarbonization rate $\mathcal{R}$, the investor takes bigger positions for assets SG and BB while reducing the exposure to assets SB and BG. However, using this framework we can see that reducing the portfolio WACI induces an increase of the tracking error risk $\sigma(\pi^*|b)$. As an example, reducing the portfolio WACI by 50% entails a tracking error volatility of 90.48 basis points, meaning that the portfolio's returns may deviate from the benchmark's return by around 0.9% on average. This may be interpreted as the risk of diverging from the benchmark, associated with decarbonizing the portfolio.
In summary, we can see that the decarbonization framework offers a simpler, more intuitive

approach for asset managers willing to reduce their carbon footprint compared to ESG-efficient frontier mentioned above.

$\mathcal{R}$	0	20%	30%	40%	50%	60%	70%	80%	$\mathcal{CI}$
π^*_{SB}	7.59	7.45	7.38	7.31	7.24	7.18	7.11	7.04	0.2392
π^*_{SG}	8.47	8.59	8.65	8.71	8.77	8.83	8.89	8.96	0.0028
π^*_{BB}	-7.49	-7.53	-7.55	-7.57	-7.60	-7.62	-7.64	-7.66	0.2139
π^*_{BG}	-8.38	-8.32	-8.29	-8.26	-8.23	-8.20	-8.17	-8.14	0.0036
$\sigma(\pi^*\|b)$	0	36.19	54.29	72.38	90.48	108.57	126.67	144.76	
WACI(π^*)	0.21	0.17	0.14	0.12	0.10	0.08	0.06	0.04	
$\bar{\mathcal{S}}(\pi^*)$	1.08	0.87	0.76	0.65	0.54	0.43	0.32	0.22	

Table 16: This table presents the optimal portfolio weights for decarbonizing a portfolio relative to an optimized benchmark. The benchmark is the a market cap value weighted portfolio of the 4 portfolios formed on size and carbon intensity. The quadratic optimization is solved for different carbon reduction rate ($\mathcal{R}$) targets. $\mathcal{CI}(\pi^*)$ shows the tracking error of the optimal portfolio π^* versus this benchmark. $\mathcal{CI}(\pi^*)$ represents the carbon footprint of the optimal portfolio measured in KtCO2/\$M using the weighted average carbon intensity (WACI) approach. The results illustrate the asset allocation tradeoffs between lowering carbon and minimizing active risk against the optimized benchmark.

7 Conclusion

This project demonstrated emerging quantitative methods for incorporating ESG criteria in asset allocation. We used two metrics to quantify ESG performance: carbon footprint and carbon risk. While both leverage a company's carbon intensity, they capture different aspects of ESG performance: emissions efficiency and potential financial risks associated with poor ESG performance. Furthermore, we showed two frameworks for integrating ESG data in portfolio optimization. However, the inconcistency of certain results associated with the (Pedersen et al., 2021) approach highlight highlight the need to carefully handle financial and ESG data to generate stable, trackable outcomes. This motivates the use of simpler methods including the optimization relative to a benchmark.

Using assets from the STOXX Europe TMI index, we applied the framework (Pedersen et al., 2021) and found that investors ignoring carbon data had higher excess returns than ESG-motivated investors pre-2020, but also more risk. More importantly, we saw that this trend reversed starting from 2020, where ESG investors showed superior risk-adjusted returns and ESG performance. This suggests integrating ESG criteria could improve financial and sustainability outcomes in coming years. Additionally, using the carbon risks has shown to provide extra information on the potential financial risks associated to carbon inefficient firms. This, may motivate the use of this metric as a forward looking measure to evaluate financial risks associated with investing in poor ESG performing assets.

However, some limitations exist in this thesis analysis. First, the ESG data used was limited to carbon intensity. Although, it is becoming a standard sustainability metric (Eurosif 2023; Gorgen et al., 2019; Roncailli et al., 2020), this lacks the incorporation of additional material data, especially for the Social (S) and Governance (G) pillars. In this context, further metrics can be used to evaluate companies on the ESG basis. This can include but is not limited to evaluating the happiness and social well-being of workers within a company to estimate its future performance[14].

Although more formal research papers found similar results with regards to the existence of a carbon risk premium (Gorgen et al., 2019; Faccini et al., 2023), further validation steps could be taken to improve the statistical properties of the model. These include estimating time varying exposures to the carbon risk (Roncailli et al., 2021) or using alternative measures of the carbon factor. The former can potentially better estimates, especially when considering the shift in ESG beliefs after the Paris agreement or the COVID pandemic.

Finally, some extensions could further improve the models used in this project. First, the ISSB revealed its global sustainability reporting standards[15], which should take place in January 2024. This should provide additional ESG and emission related data, especially for Scope 3 emissions which have shown a lot of uncertainty among ESG related research. Leveraging this new data into the existence models could provide further insights about the financial risks associated with unsustainable companies. Additionally, net zero investment strategies are emerging within the asset management industry[16], these require a yearly reduction of the carbon emissions associated with a portfolio of financial assets. Moreover, they require the forecast of future carbon emissions of the underlying assets. In this context, the benchmark optimization framework can be extended to solve inter-temporal asset allocation problems. In simple terms, this can be done by adding a penalty term to account for and optimize the costs related to rebalancing the portfolio over time. Overall, leveraging improved data, identifying alternative ESG risks, and incorporating forward looking estimations are fruitful areas to consider for sustainable investment strategies.

[14] https://www.robeco.com/en-int/glossary/sustainable-investing/social-sustainability

[15] https://www.ey.com/en_nl/ifrs/issb-issues-ifrs-1-sustainability-related-disclosures

[16] https://www.iigcc.org/net-zero-engagement-initiative

References

[1] Avramov, Doron and Cheng, Si and Lioui, Abraham and Tarelli, Andrea, Sustainable Investing with ESG Rating Uncertainty (July 26, 2021). Journal of Financial Economics (JFE), Vol. 145, No. 2, 2022, Available at SSRN: https://ssrn.com/abstract=3711218 or http://dx.doi.org/10.2139/ssrn.3711218

[2] Bolton, Patrick and Kacperczyk, Marcin T., Do Investors Care about Carbon Risk? (October 30, 2020). Columbia Business School Research Paper Forthcoming, Journal of Financial Economics (JFE), Forthcoming, European Corporate Governance Institute – Finance Working Paper 711/2020, Available at SSRN: https://ssrn.com/abstract=3398441 or http://dx.doi.org/10.2139/ssrn.3398441

[3] Bolton, Patrick and Kacperczyk, Marcin T., Global Pricing of Carbon-Transition Risk (August 5, 2022). Journal of Finance, Forthcoming, Available at SSRN: https://ssrn.com/abstract=3550233 or http://dx.doi.org/10.2139/ssrn.3550233

[4] Cheema-Fox, Alexander and LaPerla, Bridget Realmuto and Serafeim, George and Turkington, David and Wang, Hui, Decarbonizing Everything (January 28, 2021). Financial Analysts Journal, 2021, 77(3), 93-108. DOI: 10.1080/0015198X.2021.1909943., Available at SSRN: https://ssrn.com/abstract=3693941 or http://dx.doi.org/10.2139/ssrn.3693941

[5] Clark, Gordon L. and Feiner, Andreas and Viehs, Michael, From the Stockholder to the Stakeholder: How Sustainability Can Drive Financial Outperformance (March 5, 2015). Available at SSRN: https://ssrn.com/abstract=2508281 or http://dx.doi.org/10.2139/ssrn.2508281

[6] Faccini, Renato and Matin, Rastin and Skiadopoulos, George and Skiadopoulos, George, Dissecting Climate Risks: Are they Reflected in Stock Prices? (April 3, 2023). Available at SSRN: https://ssrn.com/abstract=3795964 or http://dx.doi.org/10.2139/ssrn.3795964

[7] Gorgen, M., Jacob, A., Nerlinger, M., Riordan, R., Rohleder, M., and Wilkens, M. (2019), Carbon Risk, SSRN, https://www.ssrn.com/abstract=2930897

[8] G. Halbritter, G. Dorfleitner, The wages of social responsibility — where are they? A critical review of ESG investing Rev. Financ. Econ., 26 (2015), pp. 25-35

[9] Hirschberger, M., Steuer, R. E., Utz, S., Wimmer, M., Qi, Y. (2013). Computing the nondominated surface in tri-criterion portfolio selection. Oper. Res., 61(1), 169e183.

[10] Hsu, Po-Hsuan and Li, Kai and Tsou, Chi-Yang, The Pollution Premium (May 10, 2022). Journal of Finance, Forthcoming , Available at SSRN: https://ssrn.com/abstract=3578215 or http://dx.doi.org/10.2139/ssrn.3578215

[11] Jagannathan R., Georgios Skoulakis, Zhenyu Wang, CHAPTER 14 - The Analysis of the Cross-Section of Security Returns, edited by YACINE AÏT-SAHALIA, LARS PETER HANSEN, Handbook of Financial Econometrics: Applications, Elsevier, Volume 2, 2010, Pages 73-134, ISSN 15684997

[12] Kotsantonis, Sakis, and George Serafeim. "Four Things No One Will Tell You About ESG Data." Journal of Applied Corporate Finance 31, no. 2 (Spring 2019): 50–58.

[13] Kwiatkowski, Denis Phillips, Peter C. B. Schmidt, Peter Shin, Yongcheol, 1992. "Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root?," Journal of Econometrics, Elsevier, vol. 54(1-3), pages 159-178.

[14] Le Guenedal, Théo and Lombard, François and Roncalli, Thierry and Sekine, Takaya, Net Zero Carbon Metrics (February 13, 2022). Available at SSRN: https://ssrn.com/abstract=4033686 or http://dx.doi.org/10.2139/ssrn.4033686

[15] Le Guenedal, Théo and Roncalli, Thierry, Portfolio Construction with Climate Risk Measures (January 3, 2022). Available at SSRN: https://ssrn.com/abstract=3999971 or http://dx.doi.org/10.2139/ssrn.3999971

[16] Lasse Heje Pedersen, Shaun Fitzgibbons, Lukasz Pomorski, Responsible investing: The ESG efficient frontier, Journal of Financial Economics, Volume 142, Issue 2, 2021,

[17] Loyson, Philippe and luijendijk, rianne and van Wijnbergen, Sweder, The pricing of climate transition risk in Europe's equity market (August 10, 2023). De Nederlandsche Bank Working Paper No. 788, Available at SSRN: https://ssrn.com/abstract=4540103 or http://dx.doi.org/10.2139/ssrn.4540103

[18] Markowitz, Harry. "Portfolio Selection." The Journal of Finance 7, no. 1 (1952): 77–91. https://doi.org/10.2307/2975974

[19] Pukthuanthong, Kuntara and Roll, Richard W. and Subrahmanyam, Avanidhar, A Protocol for Factor Identification (June 16, 2018). Forthcoming, Review of Financial Studies, Available at SSRN: https://ssrn.com/abstract=3197646

[20] Renneboog, L., Ter Horst, J., Zhang, C. (2008). Socially responsible investments: Institutional aspects, performance, and investor behavior. Journal of banking finance, 32(9), 1723-1742.

[21] Roncalli, Théo and Le Guenedal, Théo and Lepetit, Frederic and Roncalli, Thierry and Sekine, Takaya, Measuring and Managing Carbon Risk in Investment Portfolios (August 26, 2020). Available at SSRN: https://ssrn.com/abstract=3681266 or http://dx.doi.org/10.2139/ssrn.3681266

[22] Schmidt, Peter Steffen and von Arx, Urs and Schrimpf, Andreas and Wagner, Alexander F. and Ziegler, Andreas, Common Risk Factors in International Stock Markets (August 5, 2019). Financial Markets and Portfolio Management, 2019, Volume 33, Issue 3, pp 213–241, Available at SSRN: https://ssrn.com/abstract=1738315 or http://dx.doi.org/10.2139/ssrn.1738315

[23] Pukthuanthong, Kuntara and Roll, Richard W. and Subrahmanyam, Avanidhar, A Protocol for Factor Identification (June 16, 2018). Forthcoming, Review of Financial Studies, Available at SSRN: https://ssrn.com/abstract=3197646

Appendix A

Lag	Autocorrelation	D-W Statistic	p-value
1	0.099	1.802	< 0.01
2	0.124	1.751	< 0.01
3	0.059	1.882	< 0.01
4	0.033	1.933	< 0.01
5	0.024	1.952	0.056
6	0.021	1.957	0.098
7	0.000	2.000	0.964
8	-0.003	2.006	0.774
9	-0.009	2.017	0.446
10	-0.018	2.034	0.178
11	-0.020	2.040	0.100
12	-0.027	2.053	0.032

Table 17: This table shows the Durbin Watson test statistic and associated p-values to test for serial collinearity in monthly returns. The test was generated using a panel data regression of asset monthly excess returns of 35 portfolios on 4 factor returns. These include 25 portfolios formed on size and BE/ME (value) and 10 industry portfolios sorted using the GICS sector identification. The sample period ranges from July 2007 until June 2022. The global factors MKT, SMB, HML and the carbon factor BMG are constructed using European stocks from the STOXX Europe TM index.